بسم الله الرحمن الرحيم

جامعة اليرموك

كلية التربية

قسم الإرشاد وعلم النفس التربوي

معادلة درجات صور مختارة من الامتحان الوطني للغة الإنجليزية في الأردن وفق نظرية الاختبار ونظرية الفقرة

إعداد

مهدي محمد فهد عليمات

بإشراف

الأستاذ الدكتور أحمد سليمان عوده

حقل التخصص: القياس والتقويم

2013/2012

معادلة درجات صور مختارة من الامتحان الوطني للغة الإنجليزية في الأردن وفق نظرية الاختبار ونظرية الفقرة

إعداد

مهدي محمد فهد عليمات

بكالوريوس رياضيات، جامعة اليرموك، 1992

دبلوم عالي رياضيات، جامعة اليرموك، 1997

دبلوم عالي أساليب تدريس الرياضيات، جامعة اليرموك، 1999

ماجستير قياس وتقويم، جامعة اليرموك، 2010

قدمت هذه الأطروحة استكمالاً لمتطلبات الحصول على درجة الدكتوراه في تخصص القياس والتقويم من جامعة اليرموك، إربد، الأردن.

وافق عليها

أ.د. أحمد سليمان عوده .... رئيساً

أستاذ في البحث والتقويم التربوي، جامعة جدارا

أ.د ساري سليم سواقد .... عضواً

أستاذ في القياس والتقويم، جامعة مؤتة

أ.د. أمل عبدالله خصاونة .... عضواً

أستاذ في مناهج الرياضيات وأساليب تدريسها، جامعة اليرموك

أ.د. أحمد يوسف قواسمة .... عضواً

أستاذ في القياس والتقويم ، جامعة اليرموك

د. نضال كمال الشريفين.. عضواً

أستاذ مشارك في القياس والتقويم ، جامعة اليرموك

تاريخ مناقشة الرسالة: 2013/08/01

الإهـــــــداء

إلى أغلى الناس

إلى من يهمهم أمر نجاحي

إلى الذين لم ولن أشك يوماً في مكانتي في قلوبهم

والدتي الحنون التي أدين لها بكل حياتي ولن أفيها حقها مهما قدمت . . .

روح والدي رحمه الله . . .

روح أخي المرحوم أبو محمد الذي لم أعرف أباً سواه . . .

زوجتي الحبيبة التي تربعت على عرش قلبي وأسرت جوارحي وشاطرتني المشقة والعناء

أولادي الأحباء فلذة كبدي

إخوتي وأخواتي الأعزاء الذين يحملون همي . . . ويفرحون لفرحي ويحزنون لحزني ...

وإلى كل من علمني وله فضل علي

أهدي هذا العمل المتواضع

راجياً من الله تعالى أن يكون في ميزاني يوم الدين

شكر وتقدير

الحمد لله الذي جعلني من المسلمين، وأعانني بفضله على إتمام هذا العمل الذي أرجو منه -تعالى- أن يتقبله خالصاً لوجهه الكريم. وأصلي وأسلم على نبيه العربي الهاشمي الأمين وعلى آله وصحبه ومن استن بسنته إلى يوم الدين. وبعد

فبعد أن منّ الله -تعالى- علي بإتمام دراستي المتواضعة هذه، والتي أرجو من الله العليم أن تكون قد أفادت كل من له علاقة بموضوعها، سواء وزارة التعليم العالي والبحث العلمي الأردنية أو من له علاقة بالاختبارات والقياس والتقويم في هذا البلد المبارك، فلا بد لي من تقديم جزيل شكري وامتناني وحبي العظيم لأستاذي ومعلمي وملهمي العلامة الأستاذ الدكتور أحمد سليمان عوده على ما قدمه لي من نصح وإرشاد وحفاوة واهتمام و رعاية، سواء أثناء قيامي بإجراء هذه الأطروحة أو أثناء تدريسي، ولما تحمله معي وما منحني إياه من فائض علمه، فجزاه الله عني خير الجزاء ومتعه بدوام الصحة والعافية وأدامه ذخراً لهذا البلد.

كما يطيب لي أن أتقدم بجزيل الشكر والامتنان إلى أستاذي الفاضل الدكتور نضال الشريفين الذي لم يبخل علي بعلمه في هذا المجال ودراساته التي تناولت موضوعات معادلة الاختبارات، كما أتقدم بوافر الشكر لجميع أساتذة القياس والتقويم في جامعة اليرموك الذين أدين لهم بما تعلمته منهم، فجزاهم الله عني خير الجزاء، ولا بد أن لا أنسى في هذا المقام شكر السيد فادي الروسان -أمين سر اللجنة العليا للامتحان الوطني للغة الإنجليزية- لما أبداه من تعاون لتسهيل مهمتي وإجابته عن جميع التساؤلات حول الاختبار، وتزويدي بالبيانات المطلوبة للإجابة عن أسئلة الدراسة، فجزاه الله عني خير الجزاء.

وأتقدم بالشكر لزملائي وزميلاتي في الدراسة وأخص منهم الأخوة عبدالله عبيدات ، محمد مساعده، مصطفى مساعده، علي الصالح وأختي العزيزة أروى الحواري التي لم تبخل علي بالنصح. وأشكر مديري و زملائي في مدرسة المفرق الثانوية الأولى للبنين الذين كان لدعمهم وتشجيعهم الدائم لي أكبر الأثر في إتمام هذه الدراسة وأخص منهم بالذكر الأخ العزيز عمر الصباغ. وأخيراً أتقدم بشكري وتقديري وعظيم امتناني لكل من أسهم في إنجاز هذه الدراسة معنوياً أو مادياً.

فهرس المحتويات

فهرس الجداول

الملخص

عليمات، مهدي محمد. معادلة درجات صور مختارة من الامتحان الوطني للغة الإنجليزية في الأردن وفق نظرية الاختبار ونظرية الفقرة. رسالة دكتوراة، جامعة اليرموك 2013 (المشرف: أ.د. أحمد سليمان عوده).

هدفت الدراسة الحالية إلى معادلة درجات ثلاث صور مختارة من الامتحان الوطني للغة الإنجليزية في الأردن وفق نظرية الاختبار ونظرية الفقرة، ولتحقيق هدف الدراسة تم الحصول على البيانات الخاصة بنتائج تطبيق ثلاث صور للاختبار من الجهة المشرفة على الاختبار- وزارة التعليم العالي والبحث العلمي- على المفحوصين وعددهم (979، 854، 899) من أصل 7042 مفحوص.

تم فحص الخصائص السيكومترية لصور الثلاث(الثبات، تمييز الفقرات، وصعوبة الفقرات)، وفحص الافتراضات التي تقوم عليها النماذج المستخدمة القائمة على نظرية الفقرة. وبناءً عليه تم تحديد الفقرات والأفراد المطابقين من كل صورة من صور الاختبار لنماذج نظرية الفقرة. وتمت بعد ذلك معادلة تقديرات قدرات الأفراد المطابقين لنماذج نظرية الفقرة، والتحقق من دقة وفاعلية المعادلة باستخدام محك الصدق التقاطعي. بعد ذلك، ولكل نموذج من نماذج نظرية الفقرة، تم تكوين ثلاث صور جديدة مكونة من الفقرات المطابقة للنموذج من فقرات صور الاختبار الأصلية.

تمت إجراءات معادلة الصور الجديدة من الاختبار باستخدام المعادلة الخطية وباستخدام الرتب المئينية كطريقتين للمعادلة التي ترتكزعلى نظرية الاختبار. كما تم التحقق من فاعلية كل طريقة من طرق المعادلة بواسطة محكي الخطأ المعياري للمعادلة الخطية والمئينية، والصدق التقاطعي للمعادلة الخطية والمئينية.

وتم التوصل إلى النتائج التالية:

كانت معاملات ثبات الاتساق الداخلي في الصور الثلاث عالية حيث بلغت (0.92)، (0.91)، و(0.92). وتبين عدم وجود فروق ذات دلالة إحصائية بين متوسطات صعوبة الفقرات وتمييزها، ولدرجات المفحوصين وتبايناتها تُعزى للصورة. وأظهرت نتيجة التحليل العاملي تحقق افتراض أحادية البعد ولكل صورة من صور الاختبار الثلاث. كما وتبين عدم وجود فرق دال إحصائياً بين نسب الفقرات المطابقة لنماذج نظرية الفقرة على مستوى الصور الثلاث، بينما تبين وجود اختلاف دال إحصائياً بين نسب الأفراد المطابقين لنماذج نظرية الفقرة على مستوى الدورات الثلاث، وعدم وجود فرق دال إحصائياً بين متوسطات القدرات المقدرة وفق كل نموذج من نماذج نظرية الفقرة الثلاث. وعند مقارنة طرق المعادلة للفقرات المطابقة لنماذج نظرية الفقرة باستخدام طريقتي المعادلة التي ترتكز على نظرية الاختبار (الخطية والمئينية)، تبين أن المعادلة للصور المكونة من الفقرات المطابقة للنموذج أحادي المعلمة كانت الأكثر دقة وفاعلية، تليها معادلة الصور المكونة من الفقرات المطابقة للنموذج ثنائي المعلمة، ثم الصور المكونة من الفقرات المطابقة للنموذج ثلاثي المعلمة، وذلك سواء باستخدام محك الخطأ المعياري للمعادلة أو معامل الصدق التقاطعي. وعند استخدام محك الصدق التقاطعي للمقارنة بين فاعلية طرق المعادلة وفق النظريتين تبين أن المعادلة وفق النموذج أحادي المعلمة كانت الأكثر فاعلية تليها المعادلة وفق النموذج ثنائي المعلمة فالمعادلة وفق النموذج ثلاثي المعلمة، ثم المعادلة المئينية وأخيراً المعادلة الخطية.

(**الكلمات المفتاحية:** الامتحان الوطني، اللغة الإنجليزية، نظرية الاختبار، المعادلة، نظرية الفقرة).

الفصل الأول

خلفية الدراسة ومشكلتها

المقدمة:

تُعد اللغة الإنجليزية واحدة من أكثر لغات العالم استخداماً في مجال البحث العلمي في الوقت الحالي، لذا فإن تعلم هذه اللغة أصبح شرطاً ضرورياً لمن أراد أن يتابع دراساته العليا، ومن هنا جاءت توجهات وزارة التعليم العالي والبحث العلمي في الأردن بجعل إتقان اللغة الإنجليزية شرطاً أساسياً للتخرج ولجميع طلبة الدراسات العليا في الجامعات الأردنية، ثم جعلته شرطاً للقبول في برامج الدراسات العليا.

تعتبر الاختبارات من أدوات القياس التربوي المهمة فهي متعددة الأغراض والأهداف، كما تعد الاختبارات مؤشراً هاماً يستند إليها في عملية التقويم، فعندما نختبر الطلبة إنما نسعى لقياس مستوى التحصيل لدى التلاميذ، والوقوف على مدى التقدم الذي يحرزه الطلبة، فللاختبارات المدرسية دور أساسي في عملية التعليم، وتستخدم نتائجها في اتخاذ القرارات التربوية، حيث تستخدم علامات الطلبة لتوزيعهم على شعب متجانسة من حيث قدرتهم العقلية والتحصيلية، وفي مقارنة الطلبة بعضهم ببعض أو في تشخيص جوانب القوة والضعف لدى التلاميذ، كما تستخدم في عملية القبول والانتقاء والتوظيف والتنبؤ... وغير ذلك، فكلما زادت أهمية القرار الذي يتوقع اتخاذه، زادت الحاجة لأن تكون المعلومات التي يتم الحصول عليها من الاختبار أكثر دقة، وصلة بالغرض الذي اعد الاختبار من اجله.ومن الأمور المهمة التي يجب أن تؤخذ بعين الاعتبار أن الاختبارات هي في مجملها تطوير للعملية التعليمية، وهذا لا يتم إلا

من خلال بناء اختبار يقيس مستوى التحصيل من ناحية وأن يكون فيه ما يعمل على تقويم عملية التدريس من ناحية أخرى(عبد الهادي، 2001).

يفيد تقرير حصل عليه الباحث من وزارة التعليم العالي والبحث العلمي بأن وزارة التعليم العالي والبحث العلمي في الأردن قررت أن يكون اجتياز امتحان التوفل الدولي (اختبار اللغة الإنجليزية كلغة أجنبية (TOEFL)-Test of English as Foreign Language) والحصول على علامة النجاح (500) أو ما يعادلها من امتحانات اللغة الإنجليزية الأخرى، شرطاً للتخرج لطلبة الدراسات العليا، وأجاز القبول بعلامة (450) لبعض التخصصات الإنسانية ولكن بموافقة المجلس. وكان ذلك في جلسة مجلس التعليم العالي المنعقدة بتاريخ 2004/7/1.

وذكر التقرير أنه و بتاريخ 2011/11/21 ناقشت اللجنة العليا للامتحان الموحد موضوع إعداد امتحان وطني للغة الإنجليزية، ونسبت اللجنة إلى مجلس التعليم العالي باعتماد امتحان وطني للغة الإنجليزية كأحد الامتحانات المعتمدة لأغراض القبول في برامج الدراسات العليا ليصبح أحد الامتحانات المعتمدة بجانب الامتحانات الأخرى (التوفل ITP، التوفل IBT، وامتحان IELTS، . . . إلخ).

تم تطبيق الامتحان الوطني للغة الإنجليزية لأول مرة بتاريخ 2012/05/12 والتطبيق الثاني كان بتاريخ 2012/06/2، وأصبح بعد ذلك يعقد أسبوعياً. وجميع فقرات الاختبار هي من نوع الاختيار من متعدد، مكونة من متن يليه أربعة خيارات، واحد فقط منها صحيحة. و تؤخذ فقرات الاختبار من بنك للأسئلة معد لهذا الغرض – تم أخذ فقراته من فقرات الامتحان الموحد المذكور سابقاً–، حيث تختلف فقرات الاختبار في كل دورة يُعقد بها عن غيرها من الدورات،

وبحيث يكون لكل دورة أربعة نماذج A,B,C,D تختلف في ترتيب الفقــرات فقــط، ويــوزع المفحوصون على هذه النماذج بشكل عشوائي.

يشير علام (2005) إلى أن الاختبارات واســعة النطــاق (Large Scale Testing Programs) تتطلب في غالبية الأحيان توافر صور متعددة من الاختبارات المختلفة لــضمان سرية الاختبارات، وأن الصور المتعددة ضرورية في اختبارات القبول كالتي تستخدم في قبول انتقاء الطلبة للالتحاق بكليات معينة. وأنه ينبغي إجراء معادلة للصور المختلفة لتحويل الدرجات الخام التي يحصل عليها المفحوصين من صورتي اختبار معين بحيث تصبح الدرجات المحولة لكل منها متعادلة. ففي حالة اختبارات القبول ينبغي إجراء معادلة للصور المتعددة، لكي نتمكن من المقارنة بين الطلبة الذين يُطبق عليهم صور من الاختبار نفسه، وبدون ذلك تصبح المقارنة غير عادلة، وذلك لأنه ربما تكون فقرات إحدى صور الاختبار أكثر صعوبة من صور أخرى.

وللحفاظ على سرية الامتحان الوطني للغة الإنجليزية وخوفاً من الإخلال بها، وحتى لا يتم إعادة تطبيق الامتحان نفسه مرة أخرى وبذلك تصبح فقراته معروفة، تقوم اللجنــة الفنيــة القائمة على بناء الامتحان ببناء صورة جديدة للامتحان في كل دورة تختلف فقراتها عن فقرات صور الامتحان في الدورات السابقة بحيث تقيس نفس المهارات وبنفس الأوزان، على افتراض أن هذه الصور متكافئة تماماً، رغم عدم وجود ما يدل على تكافؤها. وحتى يمكن مقارنة نتــائج هذه الصور واتخاذ قرارات صحيحة بناءً على نتائجها لا بد من إجراء معادلــة (Equating) لهذه الصور من خلال اقتران تناظر بين صور الامتحان يتم من خلاله ربط كل درجة في إحدى الصور لما يعادلها في صورة أخرى للامتحان وهكذا لبقية الصور. فقد أثبتت دراسات سابقة أن مثل هذه الاختبارات (متعددة الصور) قد تكون غير متكافئة بالرغم من ادّعاء القــائمين عليهــا بأنها متكافئة كما هو الحال في دراسة الشريفين (2009). راجياً أن تُفيد نتائج هــذه الدراســة

الجهة القائمة على الامتحان الوطني للغة الإنجليزية في الأردن وهـــي وزارة التعلـــيم العـــالي والبحث العلمي الأردنية.

وهناك مجموعة من الطرق المستخدمة لمعادلة صور الاختبار منها مـــا يرتكـــز علـــى نظرية الاختبار (Test Theory) مثل: معادلة صور الاختبار بطريقة الوسط الحسابي (Mean Equating)، والمعادلة الخطية ((Linear Equating))، والمعادلة المئينية (Equipercetile Equating)، ومنها ما يرتكز على نظرية الفقرة (Item Theory)، مثل: معادلـــة الاختبـــار باستخدام درجات القدرة (Ability Score Equating)، ومعادلة الاختبار باستخدام الـــدرجات الحقيقية (True-Score Equating)، معادلة الاختبار بطريقة الدرجات الملاحظة (الـــدرجات الخـــام) (Observed Score Equating) (Crocker & Algina, 1986; Hamhleton, Swamithan & Rogers, 1991; Kolen & Brennan, 1995; Lord, 1980; Peterson et.al, 1989).

مشكلة الدراسة وأسئلتها :

جاءت هذه الدراسة لمعرفة الخصائص الـــسيكومترية (Psychometric Properties) لثلاث صور مختلفة من الامتحان الوطني للغة الإنجليزية في الأردن الذي تعقده وزارة التعلـــيم العالي والبحث العلمي الأردنية، ثم تقصي تكافؤ (Equivalency) هـــذه الـــصور، ومعادلـــة درجات هذه درجات الصور ووضعها على تدريج مشترك (Common Scale)، كما حاولـــت هذه الدراسة تحديد أي طرق المعادلة تحقق أكثر فاعلية وفقاً لمحكي الخطأ المعياري للمعادلـــة (Standard Error of Equating) و الصدق التقاطعي (Cross Validation). وتحديداً فإن هذه الدراسة حاولت الإجابة عن الأسئلة التالية:

أسئلة الدراسة:

السؤال الأول: ما هي الخصائص السيكومترية للصور الثلاث للامتحان الوطني للغة الإنجليزية في الأردن؟

السؤال الثاني: ما قيم القدرة المتعادلة في الصور الثلاث للامتحان الوطني للغة الإنجليزية في الأردن باستخدام نماذج نظرية الفقرة (أحادي المعلمة، وثنائي المعلمة، وثلاثي المعلمة)؟

السؤال الثالث: أي طرق المعادلة تحقق أكثر فاعلية؟

أهمية الدراسة:

تنبع أهمية الدراسة من أهمية الامتحان؛ حيث أنه يتم في ضوء نتائج المفحوصين اتخاذ قرار مهم للطلبة المتقدمين للقبول في برامج الدراسات العليا وهو قبول المفحوصين في برامج الدراسات العليا في الجامعات الأردنية. وحيث أن الامتحان جديد، ولم تُجرى دراسات على مصداقيته في قياس السمات التي يُفترض أن يقيسها (مهارات اللغة الإنجليزية)، لذا جاءت هذه الدراسة للتعرف على الخصائص السيكومترية لصور مختلفة من الامتحان الوطني للغة الانجليزية في الأردن ومعادلة درجات هذه الصور وفق طرق نظرية الاختبار و نظرية الفقرة لمحاولة تحقيق العدالة في توزيع درجات الصور المختلفة لهذه الدرجات، حيث أن تحقيق العدالة مطلب أساسي في مجال الاختبارات. إضافةً إلى إعادة النظر في نُظم هذا الامتحان من خلال نتائج هذه الدراسة علّها تعود بالنفع مستقبلاً على الجهة القائمة على تطبيق الامتحان وهي وزارة التعليم العالي والبحث العلمي الأردنية.

ولتفوق نظرية الفقرة على نظرية الاختبار في حل كثير من القضايا المعاصرة في مجال الاختبارات مثل بنوك الأسئلة، والاختبارات التكيفية وغيرها فقد رأى الباحث استخدام مخرجات

نظرية الفقرة (الفقرات المطابقة لنظرية الفقرة) في كل صورة من صـــور الامتحـــان الـــثلاث كمدخلات لمعادلة صور الامتحان وفق الطرق التي ترتكز على نظرية الاختبار، وهذا الأمر لم يتم في أي دراسة سابقة.

تعريف المصطلحات:

الامتحان الوطني للغة الانجليزية (National Test of English):

اختبار تعقده وزارة التعليم العالي والبحث العلمي الأردنية، ويُعد متطلب قبول في جميع برامج الدراسات العليا في الجامعات الأردنية، ويتضمن (100) فقرة من نـــوع الاختيـــار مـــن متعدد، وتتكون كل فقرة من متن وأربعة خيارات واحدة منها فقط صـــحيحة، وتقـــيس فقـــرات الامتحـــان مهـــارات اللغـــة الإنجليزيـــة وهـــي: الاســـتيعاب الـــسمعي (Listening Comprehension) ووزنه 20% من علامة الاختبار الكلية، الاستيعاب القرائي (Reading Comprehension) ووزنه 20% من علامة الاختبار الكلية، القواعد (Structure) ووزنـــه 25% من علامة الاختبار الكلية، التعابير المكتوبة (Written Expressions) ووزنـــه 20% من علامة الاختبار الكلية، المفردات (Vocabulary) ووزنه 15% من علامة الاختبار الكلية.

معادلة الاختبار (Test Equating) :

تحويل نظام وحدات القياس الخاص (الدرجات) بإحدى صورتي اختبار إلى نظام وحدات القياس (الدرجات) الخاص بالصورة الأخرى، بحيث تصبح القياسات المستمدة من درجات كـــل من الصورتين متعادلة بعد إجراء التحويل.

المعادلة الأفقية (Horizontal Equating):

هي المعادلة التي تكون فيها صور الاختبار الواحد معدة لمنطقة سلوك واحدة أو التوزيع نفسه للقدرة أو السمة المقيسة. وبالتالي تكون لمجتمع واحد محدد من الأفراد يتمثل فيه توزيع القدرة نفسه. وقد يكون الهدف من إجراء المعادلة تعديل الفروق الناتجة عن اختلاف مستوى الصعوبة إذا كانت الصور المختلفة متقاربة في مستوى صعوبتها أو المحافظة على سرية الاختبار إذا ما تمت المعادلة من مستوى الصعوبة نفسه وعلى مجتمعات متماثلة في توزيع أفرادها على السمة المقيسة.

النموذج أحادي المعلمة (One Parametric Model):

هو نموذج لوغاريتمي ذو معلمة واحدة (معلمة الصعوبة) يساعد في تقدير احتمالية إجابة الفرد على الفقرة إجابة صحيحة بدلالة قدرة الفرد ومعامل صعوبة الفقرة، بصرف النظر عن حجم العينة وعدد الفقرات.والمعادلة الرياضية لهذا النموذج هي:

$$P_i(\theta) = \frac{1}{1+e^{-Da(\theta-bi)}}$$

حيث $P_i(\theta)$: احتمال أن يجيب المفحوص ذو القدرة (θ) على الفقرة (i) التي صعوبتها bi إجابة صحيحة.

D: عامل التدريج (Scaling factor) وهو ثابت.

النموذج ثنائي المعلمة (Two Parametric Model):

هو نموذج لوغاريتمي ذو معلمتين (الصعوبة، والتمييز) يساعد في تقدير احتمالية إجابة الفرد على الفقرة إجابة صحيحة بدلالة قدرة الفرد ومعامل صعوبة الفقرة وتمييزها بصرف النظر عن حجم العينة وعدد الفقرات. والمعادلة الرياضية لهذا النموذج هي:

$$P_i(\theta) = \frac{1}{1+e^{-Dai(\theta-bi)}}$$

حيث $P_i(\theta)$: احتمال أن يجيب المفحوص ذو القدرة (θ) على الفقرة (i) التي صعوبتها bi وتمييزها a_i إجابة صحيحة.

D: عامل التدريج (Scaling factor) وهو ثابت.

النموذج ثلاثي المعلمة(Three Parametric Model):

هو نموذج لوغاريتمي ذو ثلاث معالم يساعد في تقدير احتمالية إجابة الفرد على الفقرة إجابة صحيحة بدلالة قدرة الفرد ومعامل صعوبة الفقرة وتمييزها وتخمينها بصرف النظر عن حجم العينة وعدد الفقرات. والمعادلة الرياضية لهذا المنحنى هي:

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{Dai(\theta-bi)}}{1+e^{Dai(\theta-bi)}}$$

حيث $P_i(\theta)$: احتمال أن يجيب المفحوص ذو القدرة (θ) على الفقرة (i) التي صعوبتها bi وتمييزها a_i، وتخمينها c_i إجابة صحيحة.

المعادلة الأفقية الخطية (Linear Horizontal Equating):

ويتم من خلالها تحويل درجات إحدى صورتي الاختبار إلى تدريج درجات الصورة الأخرى، وذلك بجعل الدرجات المعيارية المناظرة لكل منهما متساوية.

المعادلة الأفقية المئينية (Percentile Horizontal Equating):

تعتبر درجات المفحوصين متعادلة إذا تساوت رتبها المئينية بالنسبة لمجموعة معينة من الأفراد.

معايير فاعلية المعادلة (Criteria of Efficiency):

المعايير المستخدمة كوسيلة للتعبير عن دقة المعادلة وفاعليتها، وتستعمل للمقارنة بـين طرق المعادلة وتصميماتها. ومنها:

الخطأ المعياري للمعادلة (Standard Error of Equating): الانحراف المعياري لعلامـات محولة لإحدى صورتي الاختبار Y المناظرة لعلامة معينة على الصورة الأخرى X وتؤخذ من تحويل مبني على عينات مستقلة اختيرت من مجتمعها. وتختلـف قيمـه بـاختلاف التـصميم المستخدم في المعادلة.

معامل الصدق التقاطعي (Cross Validation Coefficient): متوسط مربعات الانحرافات للعلامات المتعادلة عندما يُطبق على عينات مختلفة لمعرفة مدى استقرار النتائج.

التحليل العاملي (Factor Analysis):

هو طريقة إحصائية تعمل على تلخيص لعدد من المتغيرات إلـى عـدد أقـل يعـرف بالعوامل (Factors) و كل مجموعة من المتغيرات ترتبط بعامل واحد فقـط، حيـث تـرتبط المتغيرات في العامل ارتباطاً عالياً فيما بينها وضعيفاً مع الأخرى، والتحليل العـاملي يهـدف لاستخلاص مجموعة من العوامل ترتبط بالمتغيرات الأصلية على أن تفسر أكبر نسبة ممكنة من التباين للمتغيرات الأصلية أو تقليص عدد المتغيرات في عدة عوامل.

أحادية البعد (Unidimensionality):

أحد افتراضات نظرية الفقرة، ويقصد به وجود قدرة واحـدة تفـسر أداء الفـرد فـي الاختبار. ويمكن التحقق من هذا الافتراض باستخدام أسلوب التحليل العاملي، حيث تعتمد أحادية البعد على أن نسبة الجذر الكامن للعامل الأول إلى الجذر الكامن للعامل الثاني تكون نسبة كبيرة

لا تقل عن (2)، وأن ما يفسره العامل الأول مـــن التبـــاين نـــسبة كبيـــرة. (Hambleton & Swaminathan, 1985).

محددات الدراسة :

يمكن بيان محددات الدراسة بما يلي :

1- اقتصرت عينة الدراسة على المتقدمين للامتحان في ثلاث صور وهـــي الـــصور المطبقـــة بتواريخ 2012/7/7، 2012/7/14، و 2012/7/28. حيث بينت الدراسات السابقة أن عينة المفحوصين، وعينة الفقرات تؤثر في نتائج المعادلة وخصوصاً في طرق المعادلـــة وفـــق نظرية الاختبار مثل دراسة كوك وإيجنر (Cook & Eignor, 1991) والتي أشارت نتائج الدراسة إلى اختلاف المعادلة باختلاف محتوى الاختبار، وعدم ملائمـــة طريقـــة المعادلـــة المئينية، حيث أنها عانت من نقص البيانات عند العلامات المتطرفة. ودراســـة بترســـون، وكوك وستوكنج (Peterson, Cook & Stocking, 1983) والتي أظهرت أنه إذا اختلفت الاختبارات من حيث المحتوى أو طول الاختبار فإن النموذج ثلاثي المعلمة يعطـــي ثباتـــاً واستقراراً أفضل للاختبارات المعادلة وذلك باستخدام محك الصدق التقاطعي. كما أظهـــرت دراسة كيم ورفاقه (Kim, et.al., 2008) أن حجم العينة له تأثيرات مختلفة علـــى نتـــائج المعادلة باختلاف طرق المعادلة.

2- اقتصرت الدراسة على طرق المعادلة الخطية والمئينية كطـــرق معادلـــة وفـــق نظريـــة الاختبار، وعلى طريقة المعادلة باستخدام درجات القدرة(Ability score Equating) وفق نظرية الفقرة. حيث يوجد طرق أخرى للمعادلة لم تُستخدم في الدراسة، فقـــد بيّنـــت الدراسات السابقة أنه لا يوجد طريقة معادلة محددة تحقق أفضل النتائج للمعادلـــة حيـــث

يذكر سكاجز ولزتس (Skaggs & Lissitz;1986) أن تطبيق نظرية الفقرة في معادلة الاختبارات قد حظيت باهتمام الباحثين، إلا أنه يصعب الوصول إلى استنتاجات وعمل تعميمات، حيث أجريت دراسات مختلفة، وباختبارات مختلفة، وعينات مختلفة، وطرق مختلفة، لتقييم دقة نتائج المعادلة، وكانت أهم النتائج لهذه الأبحاث هي عدم وجود طريقة معادلة محددة تحقق أفضل النتائج لمعادلة الاختبارات جميعها.

3- اقتصرت الدراسة على تصميم المجموعات العشوائية المتكافئة. حيث أن نتائج الدراسة والحسابات المستخدمة فيها تختلف باختلاف التصميم المستخدم.

4- اقتصرت عملية التحقق من فاعلية طرق المعادلة وفق طرق نظرية الفقرة على محك الصدق التقاطعي (Cross Validation). فقد أشار كولن (Kolen, 1981) إلى أنه لا يوجد محك واضح وحاسم يبين أنه الأفضل فيما يتعلق بالمقارنة بين دقة الطرق المختلفة وفاعليتها في المعادلة. فاختلاف المحك المستخدم في تحديد فاعلية طرق المعادلة قد يؤدي إلى نتائج مختلفة.

افتراضات الدراسة :

1. تفترض الدراسة أن البيانات التي تم الحصول عليها من الجهة القائمة على الامتحان قاعدة بيانات دقيقة.

2. تفترض الدراسة تكافؤ مجموعات المفحوصين، وهو الافتراض الذي تبنته الجهة القائمة على الاختبار؛ ولذا فإن الدراسة اعتمدت تصميم المجموعات العشوائية المتكافئة.

الفصل الثاني

الإطار النظري و الدراسات السابقة

الإطار النظري للدراسة:

يشير كولن (Kolen, 1981) إلى أنه عند الحديث عن المعادلة للصور المختلفة للاختبارات، توجد ثلاث قضايا مترابطة ومتداخلة يجب أخذها بعين الاعتبار وهي:

أولا: الأساس النظري الذي يتم بموجبه التناظر والمقابلة بين العلامات على الصور المختلفة، فهناك نظريتان رئيسيتان يتم من خلالهما اشتقاق طريقة المعادلة وهما نظرية الاختبار ونظرية الفقرة.

ثانيا: الحاجة إلى تصميم لجمع البيانات للمعادلة، فهناك عدة أنواع من التصاميم لجمع البيانات، حيث تتدخل عدة اعتبارات عملية في اختيار التصميم المناسب.

ثالثا:الطرق الإحصائية التي ستستخدم في تقدير العلامة المناظرة، مثل الطريقة الخطية والطريقة المئينية.

وهنا سيتم تناول شرحاً للمفاهيم الأساسية التي تقوم عليها الدراسة في إطار نظرية الاختبار ونظرية الفقرة، وتحديداً المفاهيم ذات الصلة المباشرة وغير المباشرة بمعادلة نماذج الاختبارات في حال وجود مسوغات لهذه المعادلة، وهذه المفاهيم هي التالية: معادلة الاختبار ومسوّغاتها، شروط المعادلة، أنواع المعادلة (الأفقية والعمودية)، تصميمات جمع البيانات في المعادلة، طرق المعادلة وفق النظريتين وخطواتها، والمعايير المستخدمة في الحكم على دقة وفاعلية الطرق المختلفة لمعادلة الاختبارات. وقد تم تعريف المصطلحات اصطلاحاً وإجرائياً بناءً على الإطار النظري الذي يتضمن هذه المفاهيم والموضوعات.

مفهوم معادلة الاختبار:

يتم في أغلب الاختبارات بناء نماذج مختلفة من الاختبار، وذلك لأسباب مختلفة منها: سرية الاختبار، والحد من عملية الغش في الاختبار، وتهدف معادلة الاختبار إلى تعديل الفروق في درجات المفحوصين و الناتجة عن خصائص فقرات صور الاختبارات من صعوبة، وتمييز، وتخمين بحيث تسمح باستعمال نماذج الاختبار بشكل متبادل.

يوجد نوعان من أنواع المعادلة للاختبارات: المعادلة الأفقية Horizontal Equating، والمعادلة الرأسية Vertical Equating. ففي المعادلة الأفقية تكون الصور المختلفة للاختبارات متقاربة في مستوى صعوبتها، والهدف من إجراء عملية المعادلة هو تعديل الفروق الناتجة عن هذا الاختلاف، بينما في المعادلة الرأسية تكون صور الاختبارات مختلفة في مستوى صعوبتها، وكذلك تكون عينات الأفراد التي تطبق عليها هذه الصورة مختلفة في مستوى قدرتها (Skaggs, Lissitz, 1986).

ويمكن تعريف معادلة الاختبار (Test Equating) على أنها تحويل نظام وحدات القياس الخاص (الدرجات) بإحدى صورتي اختبار إلى نظام وحدات القياس (الدرجات) الخاص بالصورة الأخرى، بحيث تصبح القياسات المستمدة من درجات كل من الصورتين متكافئة بعد إجراء التحويل.(الشريفين، 2003).

شروط المعادلة:

اتفق المختصون على مجموعة من الشروط الضرورية لإجراء عملية معادلة الاختبارات (Hambleton & Swaminathan, 1985; Lord, 1980) وهي :

1. قياس نفس السمة (القدرة): أي أن الاختبارات المراد معادلتها يجب أن تقيس الخاصـــية نفسها.

2. العدالة أو المساواة (Equity) : هو أن تكون خصائص التوزيع التكراري المــشروط للدرجات عند مستوى معين من مستويات القدرة (θ) للاختبار (Y) بعد تحويل الدرجات هي نفسها للتوزيع التكراري المشروط للدرجات عند المستوى نفسه مــن القــدرة (θ) للاختبار (X)؛ أي $f(Y/\theta) = f(X/\theta)$.

3. اللاتباين في مجتمع الدراسة (Invariance) : أي أن تحويل الدرجات يجب أن يبقى كما هو بصرف النظر عن (عينة المفحوصين).

4. التماثل (Symmetry) : أي أن تحويل الدرجات من صورة إلى أخرى قابل للانعكــاس؛ أي أن الدالة التي تحول بها الدرجات من الصورة الأولى إلى الصورة الثانية من الاختبار و الدالة التي تحول بها الدرجات من الصورة الثانية إلى الصورة الأولى مــن الاختبـــار تعطيان نفس النتائج.

5. الثبات (Reliability) : الاختبارات التي تتمتع بثبات عالي هي التي يمكن معادلتها.

أنواع المعادلة:

يبين (Hambleton & Swaminathan, 1985) نوعين رئيسين من أنــواع معادلـــة الاختبارات وهما:

1. المعادلة الأفقية (Horizontal Equating) عندما تكون الصور المختلفة للاختبار متقاربة في مستوى صعوبتها و للمفحوصين نفس توزيع القدرات. والهدف من إجراء عملية المعادلة هو تعديل الفروق الناتجة عن هذا الاختلاف.

2. المعادلة الرأسية (Vertical Equating) عندما تكون فيه صور الاختبار مختلفة في مستوى صعوبتها وتكون توزيعات قدرات المفحوصين مختلفة. والهدف من إجراء عملية المعادلة هو تعديل الفروق الناتجة عن هذا الاختلاف.

وفي هذه الدراسة تم استخدام المعادلة الأفقية نظراً لتقارب مستوى صعوبة صور الاختبار.

تصاميم جمع البيانات في معادلة الاختبارات :

يشير كل من كروكر وألجينا (Crocker & Algina, 1986)، وهامبلتون وسواميناثان(Hambletosn & Swaminathan, 1985) إلى أنه يوجد العديد من التصميمات لجمع البيانات في معادلة الاختبارات منها: تصميم المجموعة الواحدة (Single-group Design)، وتصميم المجموعات المتكافئة أو العشوائية (Equivalent-group Design or Random Group Design)، وتصميم المجموعات العشوائية المتوازنة (Counterbalanced Random Groups Design)، وتصميم اختبار الجذع المشترك (Anchor-test- group Design)، وتصميم المجموعات غير المتكافئة ذات الاختبار المشترك(Common-Item Nonequivalent-Group- Design). وكل من هذه التصميمات له افتراضات يجب تحققها في البيانات المستخدمة، وفي الدراسة الحالية، تم اعتماد تصميم المجموعات العشوائية أو المتكافئة؛ إذ تقوم اللجان المشرفة على تطبيق هذا الامتحان بتطبيق

هذه الطريقة على المفحوصين، أي أن هذه اللجان تقوم بتوزيع المفحوصين بطريقة عشوائية إلى مجموعات، ثم توزع عليهم صور مختلفة من الاختبار الواحد، كما أن حجوم عينات المفحوصين كان كبيراً نوعاً ما. مما يؤكد صحة اختيار هذا التصميم. وفيما يلي توضيح لهذا التصميم.

تصميم المجموعات المتكافئة أو العشوائية

(Equivalent-groups Design or Random Groups Design).

ويعد من أبسط التصميمات، حيث يتم فيها تقسيم المفحوصين إلى مجموعـات بطريقـة عشوائية، ومن ثم يتم تطبيق صور الاختبار على كل مجموعة على حدة. ومن الضروري أن تكون المجموعتان متماثلتين في القدرة المقيسة بأقصى درجة ممكنة وإلا فـإن عامـل التحيـز سيتدخل في عملية المعادلة، وهذا الأمر من الصعب تحقيقه بشكل تام، ولكن يمكن تحقيقه نسبياً؛ وذلك عندما تكون حجوم عينات المفحوصين كبيرة. (Livingston, 2004).

طرق معادلة الاختبارات

أولا : الطرق التي ترتكز إلى نظرية الاختبار :

تقوم عملية المعادلة في هذه الطريقة على تحويل من علامات خام إلى علامـات خـام raw-to-raw transformation؛ بمعنى أنه يُقدر درجات خام لصورة للاختبار إلى درجـات خام لصورة آخر لنفس الاختبار.(Kolen & Brennan, 1995) وفيما يلي توضـيح لطـرق المعادلة التي ترتكز لهذه النظرية:

1. المعادلة الخطية (Linear Equating):

يتم في هذه الطريقة تحويل الدرجات رياضياً لتحقيق تـساوي المتوسـطات الحـسابية والانحرافات المعيارية للفقرات في صورتي الاختبار المراد معادلتهما.

وتعتبر الدرجتان متعادلتان على صورتي الاختبار إذا كانتا تناظران العلامة المعيارية نفسها في مجموعة معينة من المفحوصين. وطرق المعادلة الخطية جميعها تنتج تحـويلاً علـى الصورة التالية:

$$\ell_x(Y) = AY + B$$

حيث ℓ_x: العلامة المحولة. Y: علامة الاختبار التي يطبق عليها التحويل.

A: معامل التحويل. B: ثابت التحويل.

وتختلف طريقة حساب قيم A , B باختلاف التصميم المستخدم، حيث يتم حسابهما في تـصميم المجموعات المتكافئة وفقاً للعلاقتين:

$$B = \bar{Y}_{eq} - \bar{Y}_x\left(\frac{\sigma_{eq}}{\sigma_x}\right) \quad ، \quad A = \left(\frac{\sigma_{eq}}{\sigma_x}\right)$$

حيث تصبح المعادلة بالصورة النهائية على الشكل التالي:

$$Y_{eq} = \frac{\sigma_{eq}}{\sigma_x}\left(Y_x - \bar{Y}_x\right) + \bar{Y}_{eq}$$

حيث:

Y_{eq}: درجات الأفراد في نموذج الاختبار المعادل إلى درجـات الأفـراد فـي نمـوذج الاختبار X.

σ_{eq}: الانحراف المعياري لدرجات الأفراد في نموذج الاختبار المعادل.

σ_x: الانحراف المعياري لدرجات الأفراد في نموذج الاختبارX.

Y_x: درجات الأفراد في نموذج الاختبار X.

$\bar{Y}_{eq}$: متوسط درجات الأفراد في نموذج الاختبار المعادل.

$\bar{Y}_x$: متوسط درجات الأفراد في نموذج الاختبارX.

2. معادلة الاختبارات بواسطة الرتب المئينية المتساوية (Equipercenetile Equating):

نشأت فكرة هذا الأسلوب من تعريف الدرجات المتكافئة (Equivalent Scores) الذي قدمه لورد (Lord, 1950) وفلاناجان (Flanagan, 1951)، وهو أن الدرجتين اللتين يحصل عليهما فرد في صورتي الاختبار X ,Y (حيث X, Y تقيسان سمة معينة بـنفس القـدر مـن الثبات)، يمكن اعتبارهما متكافئتين إذا كانت الرتبتان المئينيتان المناظرتان لهما بالنسبة لمجموعة معينة من الأفراد متساويتان.(علام، 2005).

و تتلخص خطوات المعادلة بهذه الطريقة:

تنظم الدرجات الخام لصورتي الاختبار المراد معادلتهما من خـلال جـدول تكـراري تراكمي. وتحدد الرتب المئينية لتوزيع الدرجات الخام لكل من صورتي الاختبـار. ثـم يُرسـم المنحنى التكراري التراكمي النسبي لدرجات الاختبارين على نفس المستوى الديكارتي. وبـذلك تعتبر الدرجتين المناظرتين لرتبة مئينية معينة، متكافئتين. وعندها يمكن رسم منحنـى ممهـد (Smooth Curve) للعلامتين المتكافئتين بحيث تمثل علامات إحدى الصورتين (Y) مثلاً على المحور الأفقي وعلامات الصورة الثانية (X) على المحور العمودي.

وعادةً ما يكون شكل منحنى التحويل المئيني بين العلامات على صورتي الاختبار Y, X خطياً في حالة التماثل التام في الصورتين، أما إذا كان شكل المنحنى مقعراً للأعلى فهذا يعني أن صورة الاختبار X أصعب من صورة الاختبار Y وإذا كان المنحنى مقعراً للأسفل فهذا يعنـي أن صورة الاختبار X أسهل من صورة الاختبار Y. وإذا كان شكل المنحنى على شكل حرف S فهذا يعني أن توزيع العلامات على صورة الاختبـار X أكثـر تـسطحاً (flatter , more platykurtic) من توزيع العلامات على الصورة Y، وإذا كان توزيع العلامات على صورتي

الاختبار متماثلاً فإن منحنى التحويل سيكون خطياً. (وهذا عندما يكون المحور الأفقي ممــثلاً لعلامات الصورة Y والمحور العمودي ممثلاً لعلامات الصورة X).(الشريفين، 2003).

ومما يؤخذ على طريقة المعادلة المئينية أنها تعاني من أثر كل مــن ســقف وأرضــية الاختبار، إذ أنه بالتعريف تكون أكبر درجة ملاحظة على أحد الاختبارين معادلة لأكبر درجــة ملاحظة على د الاختبار الآخر، وهذا لا يحدث في المعادلة الخطية.

ثانيا : الطرق التي ترتكز إلى نظرية الفقرة :

يرى بيكر و القرني (Baker & Alkarni, 1991) أن من الإســهامات الكبيــرة لهــذه النظرية في الممارسة التربوية، قدرتها على وضع عدة اختبارات ومجموعات من المفحوصــين على تدريج مشترك (Common scale) في عملية القياس، وإمكانية استخدامها في المعادلــة الأفقية والرأسية للاختبار، كما أنها تحل الكثير من المشكلات التي عجزت عن حلهــا نظريــة الاختبار بشرط أن يكون النموذج المستخدم في نظرية الفقرة (IT Model) هذه مطابقاً للبيانات المعدة لعملية المعادلة (Hambleton & Swaminathan &Rogers; 1991).

وقد انبثقت عن هذه النظرية مجموعة من النماذج والتي تعرف بنماذج السمات الكامنــة (Latent Trait Models) وتهدف جميعها إلى تحديد علاقة بين أداء الفرد في الاختبار وهو ما يمكن ملاحظته مباشرة، وبين السمات أو القدرات التي تكمن وراء هذا الأداء وتفسره(الشريفين، 2003).

و يشير هامبلتون وسومنيثان (Hambleton & Swaminathan, 1985) إلى ثــلاث مزايا لهذه النظرية تتلخص في أن تقدير قدرات الأفراد متحرر من الفقرات (Item-Free). وأن معالم الفقرات لا تتأثر بالعينة المختارة من المجتمــع (Person-Free). وأن هنــاك إمكانيــة

لحساب الخطأ المعياري لتقدير درجة الدقة في قياس القدرة لكل فرد. وهناك ثلاث طرق لمعادلة الاختبار بهذه النظرية هي:(Hambleton, Swaminthan &Rogers,1991)

أ – معادلة الاختبار باستخدام الدرجة الحقيقية. True- Score Equating:

في نظرية الاختبار تعرف "الدرجة الحقيقية" بأنها الدرجة المتوقعة للمفحوص (g) على الفقرة (i).بينما في نظرية الفقرة، يعبر عنها رياضيا على الصورة التالية:

$$\sum_{i=1}^{n} p_i(\theta)$$

فلو افترضنا أن الاختبار بصورتيه يقيس السمة نفسها، وأن معلمات الفقرات للصورتين أو الاختبارين وضعتا على التدريج نفسه، فإن الدرجة الحقيقية الصحيحة للطالب هي ξ في الاختبار X، وتكون η في الاختبار Y، وتكونان مرتبطتين بدرجات القدرة بواسطة منحنى خاصية الاختبار Test characteristic curve ،بالصورة الرياضية التالية(Lord,1980):

$$\eta = \sum_{i=1}^{n} p_i(\theta) \quad , \quad \xi = \sum_{i=1}^{n} p_i(\theta)$$

ويجب أن تكون تدريجات القياس للدرجات ξ ، η مستقلة عن مجموعة المفحوصين n في الاختبار، وأن يكون تدريج قياس درجات القدرة مستقلا عن عدد فقرات الاختبارين، كما يشترط أن تكون مستويات صعوبة الفقرات متطابقة في الاختبارين.

ويؤخذ على معادلة الاختبار بهذه الطريقة، أن الدرجات الحقيقية المقدرة لا تناظر الدرجات الخام مناظرة واحد لواحد. كما أن الدرجات الحقيقية المقدرة توضع على التدريج نفسه مع الدرجات الخام (Peterson, et. Al., 1989).

ب – معادلة الاختبار باستخدام درجات القدرة. Ability score Equating :

حيث يفترض أن كلا الاختبارين يقيسان القدرة أو السمة نفسها، فإذا قسنا معلمات القـــدرة هذه، ومعلمات الفقرات في الوقت نفسه لكل من الاختبارين، تكون هذه المعلمات على التـــدريج نفسه، لذلك نحتاج إلى معادلة درجات القدرة أثناء معايرة الاختبارات، أي تقدير معلمة الصعوبة والتمييز للفقرات، وينتج عن ذلك علاقة خطية بين تدريج القدرة وتقدير معلمات الفقرات، بشرط أن يكون التقديران منفصلين.

ج – معادلة الاختبار بطريقة الدرجات المشاهدة (الخام). Observed Score Equating

تقوم فكرة معادلة الاختبار بطريقة العلامات الخام، على فكرة التنبؤ بـــالتوزيع النظـــري للعلامات الخام للاختبار عن طريق بناء التوزيع التكـــراري للعلامـــات الخـــام x للاختبـــار X لمفحوص قدرته ثيتا (θ). ومن المشكلات التي تواجه المختصين في معادلـــة الاختبـــار بهـــذه الطريقة،أنه لا تنتج عنها علامات معادلة للمفحوص الذي علامته الخـــام أدنـــى مـــن مـــستوى الصدفة،ففي العلامات الخام تكون أدنى علامة هي الصفر،بينما في العلامات الحقيقية تكون أدنى علامة هي علامة التخمين.

وفي هذه الدراسة تم استخدام طريقة معادلة الاختبار باستخدام درجات القدرة (Ability score Equating).

و تشيـــر العديـــد من البحوث التي أجريت على معادلة الاختبار بواسطة هذه النظريـــة إلى الفوائد العديدة من أهمها أنها تسمح بمعادلة الاختبار على مستوى الفقـــرة الواحـــدة، أيـــضاً تسمح باستخدام الصورة القديمة للاختبار لمعادلة درجاتها بالصورة الجديدة للاختبار متى ما تـــم وضع الدرجات في كلا الاختباريـــن، وكذلك تقديـــرات معالـــم القـــدرة على تدريـــج واحـــــد (Kolen & Brennan, 1995).

أما بالنسبة لمعادلة الاختبارات بواسطة نظرية الفقرة، فيشير كل من كـوك وايجنـور (Cook & Eignor, 1991) إلى أن الخطوات الضرورية لمعادلة الاختبارات هي:

– اختيار التصميم المناسب لجمع البيانات لمعادلة الاختبار مع الأخذ بعين الاعتبار خـصائص مجموعة المفحوصين وطبيعة الاختبارات المراد معادلتها.

– اختيار النموذج المناسب الذي يطابق التصميم المناسب.

– بناء تدريج مشترك يربط العلاقة بين السمة المراد قياسها ومعالم الفقرة.

– اختيار التدريج المناسب لوضع درجات الاختبار (أي بشكل درجة حقيقية مقدرة أو درجـة مشاهدة أو درجة قدرة)، وتقوم بهذه المهمة المعقدة رياضيا بعض البـرامج الحاسـوبية مثـل Bilog، Multilog، Logist وغيرها.

ويمكن توضيح الخطوات السابقة كما يلي :

1– اختيار التصميم المناسب :

توجد ثلاثة تصميمات تستخدمها هذه النظرية لمعادلـة الاختبـارات، وهـي تـصميم المجموعة الواحدة، تصميم المجموعات العشوائية، والتصميم ذو الاختبار المشترك، ويُعتمد حجم العينة الملائم لإجراء معادلة الاختبار بشكل صحيح للحصول على تقـديرات مـستقرة للمعـالم المستخدمة في النموذج المختار لوضع الدرجات ومعالم القدرة (السمة) على تدريج واحد، وعادة عندما يكون عدد المفحوصين كبيراً وأداة القياس طويلة يفضل التصميم الأول، في حين إذا كانت أداة القياس قصيرة وعدد المفحوصين قليل يفضل التصميم الثاني، وان كانـت هنـاك فقـرات مشتركة فيجب أن تعكس هذه الفقرات المحتوى والخصائص الإحصائية لـصورتي الاختبـار،

إضافة أن أي تصميم من التصميمات المشار إليها يرتبط بنوعية البرنامج الحاسوبي المستخدم لمعادلة الاختبار.(الدوسري، 2001).

2- وضع تقديرات المعالم على تدريج مشترك :

ويتم هنا تقدير معالم الفقرة (الصعوبة والتمييز) لكل مجموعة على حده باستخدام منحنيات خاصية الفقرة (ICC)، التي يُفترض أنها مستقلة عن عينات الأفراد الذين استمدت منهم هذه المنحنيات بشرط أن يُطابق النموذج المستخدم للبيانات، فإنه يتوقع أن تكون تقديرات كل من مجموعتي معالم الفقرات متساوية، ولا تختلف إلا في حدود قدر معين من أخطاء المعاينات، وذلك بعد تطبيق أداة الدراسة على كل مجموعة من المفحوصين، وهنا يجب إجراء عملية تحويل رياضي تضع الدرجات وتقديرات معلمات السمة على نفس التدريج المستهدف بحيث يكون الوسط الحسابي والانحراف المعياري لتوزيع مستويات صعوبة الفقرات التي تم تقديرها في المعايرة الثانية مساويا لنظيرهما اللذين تم تقديرهما في المعايرة الأولى للدرجات، ويتم تحديد نقطة الأصل ووحدة قياس القدرة تبعاً لقدرة مجموعة الأفراد المفحوصين التي استخدمت في تعيير الفقرات، وعادةً يُجعل متوسط تقديرات القدرة مساوياً صفر، والانحراف المعياري لهذه التقديرات مساوياً الواحد الصحيح، لهذا توضع تقديرات المعالم على تدريج يحدده توزيع القدرات لدى الأفراد المفحوصين.

ويمكن الحصول على تقدير معالم المنحنيات المميزة للفقرات باستخدام برامج حاسوبية تستخدم طريقة الأرجحية القصوى في الحصول على تقديرات متزامنة لمعلمة صعوبة الفقرة، ومعلم قدرة الأفراد، وتقديرات القدرة التي يتم الحصول عليها لأفراد كل من مجموعتين يمكن وضعهما على تدريج مشترك.(الحواري، 2008).

3- معادلة درجات الاختبار :

تعتبر عملية المعادلة منتهية إذا تم تقدير المعلمات لفقرات كل من صــورتي الاختبــار المستهدف، وتم وضعها في تدريج مشترك، وتم كذلك الحصول على تقدير للسمة المراد قياسها لدى المفحوصين، بحيث تكون هي نفسها في أي من صورتي الاختبار، ويؤخذ خطأ القياس بعين الاعتبار(ويكون ذلك بالاعتماد على البرامج الحاسوبية المخصصة).(الدوسري، 2001).

خطأ معادلة صور الاختبار: Equating Error

من المشكلات المهمة التي يتصدى لها المتخصصون في معادلة الاختبــارات، مــشكلة الخطأ الناجم عن عملية المعادلة نفسها، وكيف يمكن العمل على تخفيض هذا الخطأ إلــى حــده الأدنى، لرفع مستوى دقة القياس إلى الحد الأقصى، وللحصول على معادلة اختبــار صــحيحة، يمكن تفسير نتائجها تفسيرا علميا وعمليا صحيحا.

وتؤثر الأنواع المختلفة لأخطاء معادلة الاختبار في تفسير النتائج، من خــلال تطبيــق معادلة الاختبار، من هنا يتوجب اختيار التصميم المناسب لمعادلة الاختبار بدقة وعناية، للتقليــل من خطأ القياس لحده الأدنى، مع الأخذ بالاعتبار الاعتبارات العملية أثناء التطبيق.

وهنالك نوعان من الخطأ المؤثر في عملية معادلة الاختبار، الخطأ العــشوائي والخطــأ المنتظم، أما الخطأ العشوائي فهو موجود دائما عند استخدام عينات من مجتمع الدراسة لتقــدير معلم أو أكثر، حيث يمكن خفض هذا النوع من الخطأ باستخدام عينات كبيره من المفحوصــين الذين تطبق عليهم معادلة الاختبار، ثم اختيار التصميم المناسب لعملية المعادلة، فعندما يــضطر الباحث لاستخدام عينات صغيره من المفحوصين، عندها يشكل الخطأ العشوائي مشكلة كبــرى، أما الخطأ المنتظم لمعادلة الاختبار، فينتج عادة عن عدم التزام الباحث بالافتراضات والــشروط

التي تتطلبها طريقه معينه في المعادلة، فعدم ضبط الآثار الناجمة عن عاملي التعب، والخبـرة والممارسة، في تصميم المجموعات المفردة يعتبر من المصادر الأساسية للخطأ المنتظم بحيث تنخفض دقة وصحة معادلة الاختبار. (الصمادي، 2005).

المعايير المستخدمة للحكم على دقة الطرق المختلفة لمعادلة الاختبارات وفاعليتها :

يشير كولن (Kolen, 1981) إلى أنه لا يوجد محك واضح وحاسم يبين أنه الأفـضل فيما يتعلق بالمقارنة بين دقة الطرق المختلفة وفاعليتها ، ومن هذه المعايير:

- محك الخطأ المعياري (Standard Error of Equating) : ويعبر عنه بأنـه الانحـراف المعياري للعلامات المحولة من إحدى الصورتين للأخرى من تحويل مبني على عينات اُختيرت بشكل مستقل، ويختلف الخطأ المعياري للمعادلة عادةً باختلاف قيمة العلامة. وتختلـف صـيغة معادلة حسابه باختلاف التصميم المستخدم. ويحسب الخطأ المعيـاري للمعادلـة الخطيـة مـن المعادلة، و يُحسب الخطأ المعياري للمعادلة عند استخدام تصميم المجموعات العشوائية المتكافئة من خلال المعادلة التالية (Thorndike,1982):

$$S.E_y = \sqrt{\frac{2\sigma_y^2\left(Z_x^2+2\right)}{N_T}}$$

حيث N_T: العدد الكلي للأفراد في المجموعتين اللتين أجابتا عن صورتي الاختبار X, Y.

Z_x: الدرجة المعيارية المناظرة للدرجة X.

$S.E_y$: هو الخطأ المعياري للمعادلة الخطية الأفقية للعلامة y عند العلامة Z_x.

ومن خلال المعادلة السابقة يتضح أن دقة المعادلة الخطية تبلغ أكبر ما يمكن عند الوسط الحسابي ($Z_x = 0$)، أي أن قيمة الخطأ المعياري تكون أقل ما يمكن، ويزداد هذا الخطأ بالابتعاد

عن الوسط من أي من الجهتين ويعود ذلك إلى أخطاء المعاينة التـــي تــؤثر علـــى تقــديرات المتوسطات الحسابية والانحرافات المعيارية المحــسوبة للنمـــاذج المختلفــة (Thorandike, 1982).

أما في حالة المعادلة المئينية يحسب الخطأ المعياري من المعادلة التالية:

$$S.E = \sqrt{\left(\sigma_x^2 \frac{pq}{\phi^2}\right)\left(\frac{1}{N_x} + \frac{1}{N_y}\right)}$$

حيث أن P: نسبة العلامات التي تقل عن قيمة معينة في الصورة Y، (q = 1 – p)

N_x: حجم العينة للاختبار X. N_y: حجم العينة للاختبار Y.

σ_x^2: تباين العلامات على الاختبار X.

ϕ: الاحداثي الصادي لمنحنى التوزيع الطبيعي المعياري عند النقطـــة التـــي تكـــون المساحة تحتها p.

– محك الصدق التقاطعي (Cross Validation): وهـــو متوســـط مربعـــات الانحرافـــات للعلامات المتعادلة عندما يطبق على عينات عشوائية مستقلة، ويكشف عن مدى استقرار النتائج المشتقة من عمل معادلة على عينات متماثلة عند تطبيق الطرق على عينات مستقلة (Kolen & Whitney, 1982).

ويحسب معامل الصدق التقاطعي باستخدام المعادلة التالية:

$$C = \frac{\sum (Xi - Y_i)^2}{NK}$$

حيث أن C: معامل الصدق التقاطعي.

X_i: العلامات المعادلة في العينة الأولى. Y_i: العلامات المعادلة في العينة الثانية.

N: عدد العلامات الملاحظة في توزيع الصدق التقاطعي.

K: عدد الفقرات في الاختبار المتعادل (المرجع).

والمثال التالي يبين كيفية حساب معامل الصدق التقاطعي:

ليكن لدينا الاختبار X المتعادل والمكون من (15) فقرة، ونريد معادلة الاختبار Y مع الاختبار X وفق الطريقة الخطية، نأخذ عينة عشوائية من درجات المفحوصين على الاختبار X وليكن حجمها (10) درجات، ونأخذ عينتين مستقلتين (Y1,Y2) من درجات المفحوصين على الاختبار Y حجم كل منها (10) مفحوصين، وعلى سبيل المثال كانت العينات ممثلةً بالجــدول (1).

الجدول (1)
توزيع درجات عينات مثال الصدق التقاطعي

عينة X	2	5	4	7	3	8	5	2	7	9
العينة Y1	3	8	5	4	6	7	3	5	9	7
العينة Y2	4	6	8	4	2	6	3	7	6	5

فلو أردنا حساب معامل الصدق التقاطعي لمعادلة درجات الاختبارين X و Y نحــسب الوسط الحسابي والانحراف المعياري لكل من العينات الثلاث لإجراء عملية المعادلة الخطية، ثم نحسب قيمة معامل التحويل A، وثابت التحويل B ونجري المعادلة، فتكون نتائج الحسابات كما في الجدول (2).

الجدول (2)
المتوسط الحسابي والانحراف المعياري لعينات مثال الصدق التقاطعي

العينة	المتوسط الحسابي	الانحراف المعياري
X	5.2	2.49
Y1	5.7	2.06
Y2	5.1	1.85

لمعادلة العينة X مع Y1 و X مع Y2 تكون قيم معاملي التحويل A1, A2، وثوابت التحويل B1, B2 كما يلي: A1=(2.06/2.49)=0.83 , B1= 5.2- 5.7(2.06/2.49) = 1.40

A2=(1.85/2.49)=0.75 , B1= 5.2- 5.1(1.85/2.49) = 1.22

فتكون معادلتي التحويل هي:

$$\ell_x(Y) = 0.83Y + 1.40 \qquad \ell_x(Y) = 0.75Y + 1.22$$

ثم نحسب القيم المتعادلة من عينة X مع العينتين Y2 , Y1 ونحسب مجموع مربعات الانحرافات للعلامات المتعادلة في العينتين Y2 , Y1 كما في الجدول (3).

الجدول (3)
مربعات الانحرافات للعلامات المتعادلة و مجموعها في مثال الصدق التقاطعي

$(\acute{Y}2 - \acute{Y}1)^2$	Ý2 = القيم المتعادلة من العينة Y2	Ý1 = القيم المتعادلة من العينة Y1
0.1156	2.72	3.06
0.3364	4.97	5.55
0.25	4.22	4.72
0.5476	6.47	7.21
0.1764	3.47	3.89
0.6724	7.22	8.04
0.3364	4.97	5.55
0.1156	2.72	3.06
0.5476	6.47	7.21
0.81	7.97	8.87
المجموع = 3.908		

فيكون معامل الصدق التقاطعي:

.C = 3.908/(15)(10) = 0.026

الدراسات السابقة

يتناول الباحث هنا بعض الدراسات السابقة المنشورة والتي قام الباحث بالرجوع إليهـا حول موضوع معادلة الاختبار في نظرية الاختبار، وفي نظرية الفقرة على حـد سـواء. وقـد صنفت الدراسات السابقة المتعلقة بموضوع الدراسة التي توافرت لدى الباحث إلى مجموعتين:

1. الدراسات التي قارنت بين طرق المعادلة. وكان الغرض من مراجعة هـذه الدراسـات معرفة أفضل طرق المعادلة – إن كان هنالك طريقة مفضلة على غيرها– .
2. الدراسات التي تناولت موضوع معادلة درجات اختبارات محددة. أما الهدف من دراسة ومراجعة هذه الدراسات فكان التعرف على كيفية التعامل مع عمليـة المعادلـة بـشكل تطبيقي على أمثلة واقعية في معادلة الاختبارات بعد التعرف والتعمـق النظـري فـي المعادلة.

وتالياً تفصيلاً لهذه الدراسات:

أولاً : الدراسات التي قارنت بين طرق المعادلة

في دراسة أجراها (الشريفين، 2003) هدفت إلى الكشف عن مدى تحقق معايير الفاعلية في معادلة اختبارين أحدهما ثنائي التدريج والآخر متعدد التدريج وفق نماذج نظريـة الاختبـار ونظرية الفقرة. فقد استخدم الباحث معيارين للحكم على فاعلية طرق المعادلة المستخدمة همـا: معامل الصدق التقاطعي، والخطأ المعياري للمعادلة. وأشارت نتائج المعادلة الأفقية التي أجريت وفق تصميم المجموعات المتكافئة أن النموذج أحادي المعلمة كان الأكثر فاعلية مـن طريقتـي المعادلة الخطية والمئينية ؛ إذ بلغت قيم معاملات الصدق التقاطعي أقل ما يمكن للنموذج أحادي

المعلمة حيث بلغ (0.00157) ثم تزداد في المعادلة الخطية حيث بلغت قيمته (0.0075)، وتصبح أكبر ما يمكن في المعادلة المئينية حيث بلغت قيمته (0.0234). وكانت جميعها ذات فاعلية كبيرة. أما وفق معيار الخطأ المعياري للمعادلة فقد كانت المعادلة الخطية أكثر فاعلية من المعادلة المئينية، ولم يتمكن الباحث من استخدام هذا المحك في نظرية الفقرة نظراً لعدم توافر البرامج الإحصائية المناسبة.

أما دراسة جلواكي (Glowacki, 1991) والتي استخدم فيها طرق منبثقة من نظرية الفقرة، ومقارنتها بطرق نظرية الاختبار في المعادلة للتعرف على أكثرها ملاءمة لاختبار ولاية ألباما لتخريج طلبة المرحلة الثانوية، فقد توصلت إلى أن طرق نظرية الاختبار التي استخدمت أكثر ملاءمة إذا كانت صور الاختبار المراد معادلتها قد تم تطويرها بواسطة الطرق التقليدية في بناء الاختبار.

وفي دراسة هاريس وكولن (Harris & Kolen, 1990) والتي هدفت إلى معرفة مدى فاعلية طرق المعادلة باختلاف مستويات القدرة، وكانت طرق المعادلة المستخدمة هي الطريقة الخطية والطريقة المئينية، والمعادلة وفق النموذج ثلاثي المعلمة في حالة نظرية الفقرة، واستخدمت لتحقيق هذا الهدف خمس صور لاختبارات المفاهيم الحسابية، (Arithmetic Concept Test (ACD)) بحيث كانت متماثلة في درجة ثباتها وخصائص الصعوبة والمحتوى. وأشارت نتائج هذه الدراسة إلى عدم تميز أي من طرق المعادلة المستخدمة في إعطاء نتائج أفضل للمعادلة.

وفي دراسة أجراها بترسون، وكوك وستوكنج (Peterson, Cook & Stocking, 1983) هدفت إلى مقارنة فاعلية طريقة المعادلة الخطية وطريقة المعادلة وفق النموذج ثلاثي

المعلمة وذلك من حيث ثبات المعادلة للاختبارات واستقرارها، وأجريت الدراسة على عينة من (2670) طالباً وطالبة، حيث استخدم اختبار الاستعداد الدراسي (Scholastic Aptitude Test (SAT))، والذي يقيس قدرات التطور اللغوي والتفكير الرياضي، ويستخدم كمقياس لقبول الطلبة في الجامعات. وقد أظهرت نتائج الدراسة أن طريقة المعادلة الخطية هي الأكثر فاعلية إذا كانت الاختبارات المستخدمة للمعادلة متماثلة (Parallel-Tests)، وأما إذا اختلفت الاختبارات من حيث المحتوى أو طول الاختبار فإن النموذج ثلاثي المعلمة يعطي ثباتاً واستقراراً أفضل من الاختبارات المعادلة وذلك باستخدام محك الصدق التقاطعي.

وفي دراسة كولن وتيني (Kolen &Whitney, 1982) والتي هدفت إلى مقارنة مدى فاعلية أربع طرق لمعادلة اختبارات التطور التربوي العام (Test of General Education Development (GED)) باستخدام المعادلة الأفقية واستخدام معيار الصدق التقاطعي (Cross-Validation) للحكم على فاعلية طرق المعادلة التالية: الطريقة الخطية، والطريقة المئينية، والطريقة باستخدام نموذج راش، والطريقة وفق النموذج ثلاثي المعلمة. وأظهرت النتائج أنه لا توجد طريقة مثلى من بين الطرق الأربعة، على الرغم مما أبدته نماذج نظرية الفقرة من تفوق على الطرق الكلاسيكية في المعادلة.

ثانياً : الدراسات التي تناولت موضوع معادلة درجات اختبارات محددة

ففي دراسة أجراها الشريفين (2009) لمعادلة درجات سبعة نماذج مختلفة لاختبار الكفاءة في اللغة الإنجليزية الذي عقدته جامعة اليرموك للعام الدراسي 2007/2006، بيّنت نتائج الدراسة: أن النماذج السبعة لم تكن متكافئة في أوساطها الحسابية وتبايناتها. ومن أجل معادلة نماذج الاختبار ذاته، فقد تم استخدام نموذج راش (Rash) كأحد نماذج نظرية الفقرة (IT)،

وذلك بعد أن تم التحقق من أحادية البعد لكل نموذج من نماذج الاختبار. كما تم استخدام المعادلة الخطية والمعادلة المئينية كطرق للمعادلة منبثقة عن النظرية الكلاسيكية في القياس. وبيّنت نتائج الدراسة: القدرة المقابلة للدرجة (25) في نموذج (A) التي يعدها المشرفون علـــى الاختبـــار كدرجة للنجاح، تقابلها درجات قدرة متفاوتة على نماذج الاختبار الأخرى تراوحت بين (0.54−) و (0.01) وحدة لوجيت، وفي النموذج (K) يقابلها درجات خام متفاوتة تراوحت بين (20.3) و (26)، كما يقابلها رتب مئينية متفاوتة على النماذج تراوحت قيمتها بين (45) و (70) تقريبـــاً. وبناءً على النتائج، أوصى الباحث باستخدام قيم القدرة المعادلة، أو الدرجات الخام المعادلـــة أو الدرجات الخام المقابلة للرتب المئينية نفسها كأساس لاتخاذ قرار النجاح أو الرسوب.

وفي دراسة أجرتها الحواري (2008) استخدمت الباحثة كل من المعادلة الخطية الأفقية والمعادلة المئينية الأفقية كطريقتين للمعادلة وفق نظرية الاختبار، ونموذج راش الأحادي المعلمة كطريقة من طرق نظرية الفقرة لمعادلة ثلاث نماذج لاختبارات الرخصة الدولية لقيادة الحاسوب في الأردن. كشفت نتائج هذه الدراسة عن أهمية معادلة درجات نماذج الاختبارات، كما أظهرت النتائج أن المعادلة وفق نموذج راش الأحادي المعلمة هي الأفضل والأكثر فاعلية فـــي معادلـــة درجات المفحوصين من خلال وضعهم في ميزان واحد مـــشترك، بالاعتمـــاد علـــى المعـــايير السابقة، ثم يليها المعادلة الخطية وأخيراً المعادلة المئينية.

وفي دراسة أجراها هانسون وكيم (Hanson & Kim, 2002) حول معادلة الاختبارات لنماذج مختلفة من اختبارات الرياضيات (ACT)، وهي من نوع الاختيار من متعدد، اســـتخدما فيها طريقة منحنى خصائص الفقرة، بالاعتماد على تصميم المجموعات العـــشوائية المتكافئـــة، وتكونت الاختبارات من (65) فقرة أعدت بصورة نموذجين مختلفين، طُبق النموذج الأول فـــي

عام 1997 على (2163) مفحوصاً، أما النموذج الثاني فقد طُبق عام 1998 على (2909) مفحوص. وكانت تقديرات المعلمات المطبقة على نماذج الاختبارات المختلفة مقدرة باستخدام برنامج MULTILOG (وهو برنامج محوسب يتم من خلاله تقدير قدرة المفحوص باستخدام منحنى خصائص الفقرة بالاعتماد على الدرجات التي يحصل عليها الأفراد على كل فقرة من فقرات الاختبار)، وقد أظهرت النتائج فاعلية هذه الطريقة.

وفي الدراسة التي قام بها كوك وإيجنر (Cook & Eignor, 1991) والتي درست المعادلة الأفقية لصور متكافئة في البيولوجيا والرياضيات والدراسات الاجتماعية من برنامج القبول بالدراسات العليا (Grand Record Examination) (GRE). أشارت نتائج الدراسة إلى اختلاف المعادلة باختلاف محتوى الاختبار، وعدم ملائمة طريقة المعادلة المئينية، حيث أنها عانت من نقص البيانات عند العلامات المتطرفة.

تعقيب على نتائج الدراسات السابقة

يلاحظ من خلال مراجعة الدراسات السابقة حول معادلة الاختبارات باستخدام نظرية الاختبار أو نظرية الفقرة الآتي:

1. بيّنت العديد من الدراسات أن المعادلة وفق طرق نظرية الفقرة تعطي نتائج أفضل في معادلة الاختبارات، وخاصة إذا تم استخدامها بطريقة صحيحة، وباستخدام البرنامج الحاسوبي المخصص كما حصل في دراسة هانسون وكيم (Hanson & Kim, 2002). هذا على الرغم من وجود دراسات بيّنت أن طرق المعادلة وفق نظرية الاختبار قد تكون أفضل في عملية المعادلة مثل دراسة جلواكي (Glowacki, 1991)، وبعضها الآخر بيّن تكافؤ طرق النظريتين في عملية معادلة الاختبارات. حسب ما

أظهرت نتائج دراسة هاريس وكولن (Harris & Kolen, 1990). وبسبب اختلاف نتائج هذه الدراسات التي قارنت بين طرق المعادلة التي ترتكز على النظريتين فقد وجد الباحث استخدام طرق التي ترتكز على النظريتين معاً، ثم محاولة مقارنة فاعلية هذه الطرق.

2. تنوع البرامج الحاسوبية التي يتم من خلالها تقدير قدرة المفحوص باستخدام منحنى خصائص الفقرة بالاعتماد على الدرجات التي يحصل عليها الأفراد على كل فقرة من فقرات الاختبار.

3. بينت الدراسات وجود عوامل تؤثر في عملية معادلة الاختبارات، مثل عينة المفحوصين، وعينة الفقرات وخصوصاً في طرق المعادلة وفق نظرية الاختبار مثل ما حدث في دراسة كوك وإيجنر (Cook & Eignor, 1991) والتي أشارت نتائجها إلى اختلاف المعادلة باختلاف محتوى الاختبار، وعدم ملاءمة طريقة المعادلة المئينية، بسبب نقص البيانات عند العلامات المتطرفة. إضافةً إلى دراسة بترسون، وكوك وستوكنج (Peterson, Cook & Stocking, 1983) والتي أظهرت أنه إذا اختلفت الاختبارات من حيث المحتوى أو طول الاختبار فإن النموذج ثلاثي المعلمة يعطي ثباتاً واستقراراً أفضل من الاختبارات المعادلة وذلك باستخدام محك الصدق التقاطعي. لذا استخدم الباحث البيانات التي طابقت نظرية الفقرة كمدخلات لنظرية الاختبار، حيث أظهرت نظرية الفقرة تفوقاً في بناء بنوك الاختبارات

الفصل الثالث

الطريقة والإجراءات

يتناول هذا الفصل وصفا لمجتمع الدراسة وطريقة اختيار العينة، كما ويتضمن وصفاً للطرق والأساليب والمعالجات الإحصائية التي تم استخدامها وذلك للإجابة عن أسئلة الدراسة.

مجتمع الدراسة وعينتها:

تكون مجتمع الدراسة من كافة المتقدمين للامتحان في دورات الاختبار منذ أول دورة ولغاية الدورة المنعقدة بتاريخ 2012/7/28، والبالغ عددهم (7042) مفحوصاً، وجميعهم من طلبة الدراسات العليا في الجامعات الأردنية والذين كان إتقان اللغة الإنجليزية شرطاً لتخرجهم، والطلبة الذين يريدون إتمام دراساتهم العليا ممن كان إتقان اللغة الإنجليزية شرطاً لقبولهم في البرامج التي يرغبون إتمام دراساتهم العليا بها (حيث توجد قوائم تحتوي أسماء المفحوصين ودرجاتهم على الاختبار). وتقوم اللجنة التنسيقية للامتحان بحصر أعداد الطلبة المتقدمين للامتحان ومتابعتها مع الجامعات ومطابقة أوراق الإجابات مع أعداد المتقدمين له وتدقيق المعلومات وتصحيح الامتحان واستخراج النتائج وتوزيعها بعد إقرارها.

أما بالنسبة لفقرات الاختبار فكما ذكر تقرير حصل عليه الباحث من وزارة التعليم العالي والبحث العلمي فهي مأخوذة من مجموعة فقرات الامتحان الموحد (بديل التوفل) الذي عُقد أكثر من (12) مرة، ويتم اختيار الفقرات من خلال اللجنة الفنية للامتحان، ثم يتم تدقيقها واختيار المناسب منها من قبل اللجنة العليا للامتحان.

وتكونت عينة الدراسة من مجموعة المفحوصين، الذين تقدموا للاختبار في الدورات المنعقدة بتواريخ: 2012/7/7 (الدورة الأولى)، 2012/7/14 (الدورة الثانية) و 2012/7/28 (الدورة الثالثة)، والفقرات التي طُبقت عليهم. وبلغ عدد المفحوصين في الدورة الأولى (979) مفحوصاً، وفي الدورة الثانية (854) مفحوصاً، وفي الدورة الثالثة (899) مفحوصاً. وبذلك يبلغ حجم العينة الكلي (2732) مفحوصاً. وكما ذُكر سابقاً فإن المتقدمين في كل دورة يتم توزيعهم عشوائياً للتقدم للامتحان على أربعة نماذج تختلف فيما بينها فقط في ترتيب الفقرات، و يبين الجدول (4) أعداد المفحوصين في كل نموذج وفي كل دورة من الدورات الثلاث.

الجدول (4)
توزيع أفراد عينة الدراسة حسب الدورة والنموذج

الدورة	عدد المفحوصين على النموذج A	عدد المفحوصين على النموذج B	عدد المفحوصين على النموذج C	عدد المفحوصين على النموذج D	المجموع
الأولى	251	245	238	245	979
الثانية	203	211	213	227	854
الثالثة	226	224	225	224	899
المجموع	**680**	**680**	**676**	**696**	**2732**

وكما ذُكر فإن هذه النماذج (A, B, C, D) تحتوي نفس الفقرات ولكن بترتيب مختلف، وقد تم تزويد الباحث بأرقام الفقرات المتناظرة في النماذج المختلفة وعلى مستوى الدورة الواحدة؛ فعلى سبيل المثال في الدورة الأولى المنعقدة بتاريخ 7/7/ 2012 كانت الفقرة (21) في نموذج A هي نفس الفقرة (34) في نموذج B وهي نفس الفقرة (32) في نموذج C و هي نفس الفقرة (43) في نموذج D ، فقام الباحث بإعادة ترتيب نتائج المفحوصين في كل دورة

وعلى كل فقرة من فقرات الاختبار، بحيث دُمجت نتائج الفقرة (34) في نمـوذج B، و نتـائج الفقرة (32) في نموذج C، ونتائج الفقرة (43) في نموذج D مع نتائج الفقرة (21) في نمـوذج A، وهكذا بالنسبة لجميع الفقرات. وبذلك فقد تم دمج نتائج المفحوصين على كـل دورة بـشكل منفصل بحيث أصبح حجم العينات معتمداً على الدورة وليس نماذج الدورة.

وبذلك أصبحت حجوم عينات المفحوصين كبيرة مما يحقق شرط تكـافؤ المجموعـات نسبياً، فقد تم اعتماد تصميم المجموعات العشوائية المتكافئة والذي يفترض أن المجموعات التي يتم معادلتها على نماذج الاختبار هي مجموعات متكافئة. إضافة إلى التقليل من أخطاء المعادلة.

التعريف بأداة الدراسة المستخدمة (الامتحان الوطني للغة الإنجليزية في الأردن):

يفيد تقرير حصل عليه الباحث من وزارة التعليم العالي والبحث العلمي وهـي الجهـة القائمة على الاختبار فأن:

الامتحان الوطني للغة الإنجليزية في الأردن يقيس مجموعة من المهارات فـي اللغـة الإنجليزية وهو مكون من (100) فقرة موزعة على مهارات اللغة كما يلي:

- الاستيعاب السمعي (Listening Comprehension) ووزنه 20% مـن علامـة الاختبار الكلية.
- الاستيعاب القرائي (Reading Comprehenssion) ووزنـه 20% مـن علامـة الاختبار الكلية.
- القواعد (Structure) ووزنه 25% من علامة الاختبار الكلية.

- التعابير المكتوبة (Written Expressions) ووزنه 20% من علامة الاختبار الكلية.

- المفردات (Vocabulary) ووزنه 15% من علامة الاختبار الكلية.

و تعقده وزارة التعليم العالي الأردنية بالتعاون مع الجامعات الأردنية. وتعتمد العلامات التالية لأغراض النجاح في الامتحان:

1. 75% لطلبة الطب وطب الأسنان والصيدلة ودكتور الصيدلة والطب البيطري والهندسة بكافة فروعها، وتخصصات اللغة الإنجليزية كافة (أدب، ترجمة، مناهج وأساليب تدريس اللغة الإنجليزية . . . إلخ).

2. 65% لطلبة باقي التخصصات العلمية و تخصص ماجستير تخصص إدارة الأعمال (MBA: Master of Business Administration).

3. 50% لطلبة التخصصات الإنسانية، والاقتصادية، و الإدارية باستثناء (MBA) وتخصصات اللغة الإنجليزية كافة.

ومدة صلاحية نتائج الامتحان سنتين اعتباراً من تاريخ ظهور نتيجة الامتحان. وفقرات الاختبار هي من نوع الاختيار من متعدد، حيث يختار المفحوص الإجابة الصحيحة من بين أربعة مموهات بعد قراءة أو سماع – في جزء الاستماع– متن السؤال. وفقرات الاختبار ثنائية التدريج التي تأخذ علامة واحد في حالة الإجابة الصحيحة، و صفر في حالة الإجابة الخاطئة.

وتقوم على إعداده لجنة فنية من أساتذة جامعيين بالتعاون مع وزارة التعليم العالي والبحث العلمي، ويتم اختيار فقراته من بنك الأسئلة الذي تم تكوينه من خلال فقرات الامتحان الموحد الذي وتم معرفة صعوبة فقراته بعد تطبيقه سابقاً على نفس المجتمع وهم طلبة الدراسات

العليا في الجامعات الأردنية بحيث تكون نماذج الاختبار متكافئة الصعوبة. ثم يتم تدقيقها واختيار المناسب منها من قبل اللجنة العليا للامتحان. ولا يُشترط نجاح المفحوص في كل المهارات وإنما يتم تقييم المفحوص بناءً على الدرجة الكلية التي يحصل عليها، وفق علامات النجاح المشار إليها سابقاً. وحيث أن الامتحان يتم تطبيقه على مجموعة متجانسة من المفحوصين (طلبة الدراسات العليا في الجامعات الأردنية) ما يدل على أن أي اختلاف في الأداء على نماذجه ناتج من اختلاف الأسئلة وليس اختلاف المفحوصين.

ونظراً لقلة المعلومات المتوفرة عن نماذج الامتحان الوطني للغة الإنجليزية في الأردن، لأن فقراته يتم التعامل معها بسرية تامة، تظهر الحاجة إلى إخضاع هذا الامتحان للتحليل التفصيلي والإجمالي وفق نظرية الاختبار ونظرية الفقرة للأسباب المشار إليها في مسوّغات مشكلة الدراسة.

الإجراءات والمعالجة الإحصائية

وللإجابة عن أسئلة الدراسة فقد تم الحصول على البيانات ونتائج المفحوصين وتبويبها ضمن جداول خاصة وذلك للصور الثلاثة، بالتعاون مع وزارة التعليم العالي والبحث العلمي التي تقوم بتطبيق هذه الاختبارات، وقد حصل الباحث على جميع البيانات التفصيلية لدرجات المفحوصين (0 ، 1) لكل فقرة ولكل مفحوص وفي كل نموذج من نماذج الاختبار والتي يحتاجها الباحث في دراسته من الوزارة – وذلك بعد إحضار كتاب تسهيل مهمة من الجامعة- وبعد الحصول على البيانات، فقد تم إجراء المعالجات الإحصائية التالية للإجابة عن أسئلة الدراسة:

للإجابة عن السؤال الأول: ما هي الخصائص السيكومترية للصور الثلاث للامتحان الوطني للغة الإنجليزية في الأردن؟ فقد تم:

1. حساب الإحصاءات الوصفية(الوسط الحسابي، والانحراف المعياري، والتواء التوزيع) للعلامات الخام لكل صورة من صور الثلاث للاختبار. وذلك للتعرف على هذه الإحصائيات وخصائص توزيع كل صورة من صور الاختبار الثلاث ومقارنة هذه الخصائص.

2. مقارنة متوسطات الدرجات الخام للمفحوصين باستخدام تحليل التباين الأحادي.

3. اختبار تجانس التباين للدرجات الملاحظة.

4. حساب معاملات الثبات لصور الاختبار الثلاث، حيث أن أحد شروط المعادلة أن تتميز نماذج الاختبارات المراد معادلتها بالقدر نفسه من الثبات.

5. حساب قيم مؤشر الخطأ المعياري للقياس للصور الثلاث، والذي يساعد في تحديد دقة القياس.

6. حساب معاملات صعوبة الفقرات لصور الاختبار الثلاث، وحساب معاملات تمييز الفقرات لصور الاختبار الثلاث، وذلك لمعرفة الخصائص السيكومترية لصور الاختبار.

7. مقارنة متوسطات معاملات الصعوبة لفقرات صور الاختبار الثلاث باستخدام تحليل التباين الأحادي. حيث أن اختلاف متوسطات معاملات الصعوبة يُعد مؤشراً على عدم تكافؤ صور الاختبار.

8. مقارنة معاملات التمييز لفقرات صور الاختبار الثلاث باستخدام تحليل التباين الأحادي. إذ أن اختلاف متوسطات معاملات التمييز يُعد مؤشراً على عدم تكافؤ صور الاختبار. إضافة إلى فحص افتراض تساوي القدرة التمييزية لفقرات الاختبار في النموذج أحادي المعلمة.

وللإجابة عن السؤال الثاني: ما قيم القدرة المتعادلة في الصور الثلاث للامتحان الوطني للغة الإنجليزية في الأردن باستخدام نماذج نظرية الفقرة (أحادي المعلمة، وثنائي المعلمة، وثلاثي المعلمة)؟ فقد تم:

1. استخدام أسلوب التحليل العاملي (Factor Analysis) لفحص افتراض أحادية البعد لكل صورة من صور الاختبار الثلاث.

2. مطابقة الاستجابات عن فقرات كل صورة من صور الاختبار الثلاث لنماذج نظرية الفقرة: حيث تم تحليل استجابات أفراد عينة الدراسة الذين أجابوا عن فقرات كل صورة من صور الاختبار الثلاث باستخدام برنامج (BILOG) المصمم لتحليل البيانات وفقاً لنماذج نظرية الفقرة. وذلك من أجل تحديد الفقرات والأفراد غير المطابقين لنماذج نظرية الفقرة.

3. مقارنة نسب الفقرات المطابقة و نسب الأفراد المطابقين لنماذج نظرية الفقرة الثلاث. لأن اختلاف هذه النسب قد يعني عدم تكافؤ صور الاختبار الثلاث.

4. تقدير قدرات الأفراد وحساب الخطأ المعياري لتقدير قدرات الأفراد في الصور الثلاث للامتحان الوطني للغة الإنجليزية باستخدام البرنامج الإحصائي (BILOG)، عن طريق مقدرات الأرجحية العظمى. وذلك من أجل إجراء عملية المعادلة إذا كان هنالك مبرر لها.

5. رسم منحنيات خصائص الفقرات المطابقة للنموذج أحادي المعلمة في الصور الثلاث، من أجل فحص افتراض انعدام التخمين.

6. معادلة قيم تقديرات قدرات الأفراد في الصورتين الثانية والثالثة إلى قيم تقديرات القدرة للأفراد في الصورة الأولى وفق نماذج نظرية الفقرة إذا كان هنالك مبرر لعملية المعادلة باستخدام المعادلة التالية المشار إليها في رايت وماسترز (Wright & Masters, 1982):

$$\theta_{eq} = \frac{\sigma_{eq}}{\sigma_A}\left(\theta_A - \overline{\theta_A}\right) + \overline{\theta_{eq}}$$

حيث:

θ_{eq}: قدرات الأفراد في نموذج الاختبار المعادل إلى درجات الأفراد في نموذج الاختبار A.

σ_{eq}: الانحراف المعياري المعدّل من الخطأ لتقديرات القدرة للأفراد في نموذج الاختبار المعادل. ويتم حسابه من خلال المعادلة:

$$\sigma_{eq} = \sqrt{\sigma_p^2 - MSE_p}$$

حيث:

MSE_p: متوسط مربعات الأخطاء المعيارية في قياس قدرات الأفراد (Mean Square Measurement Error) ويحسب من العلاقة التالية:

$$MSE_p = \frac{\sum_{n=1}^{N} S_n^2}{N}$$

حيث:

S_n: الخطأ المعياري في تقدير قدرات الأفراد، N: عدد الأفراد

σ_A: الانحراف المعياري المعدّل من الخطأ لتقديرات القدرة للأفراد في نموذج الاختبار A.

θ_A: تقديرات القدرة في نموذج الاختبار A.

$\overline{\theta}_{eq}$: متوسط تقديرات القدرة للأفراد في نموذج الاختبار المعادل.

$\overline{\theta}_A$: متوسط تقديرات القدرة للأفراد في نموذج الاختبار A.

8. حساب معامل الصدق التقاطعي لدقة المعادلة وفق نماذج الاستجابة للفقرة.

وللإجابة عن السؤال الثالث: أي طرق المعادلة تحقق أكثر فاعلية؟ فقد تم:

1. معادلة الصور الجديدة للاختبار، وهي مكونة من الفقرات المطابقة لنماذج نظرية الفقرة في الصور الأصلية للاختبار باستخدام طرق المعادلة التي ترتكز على نظرية الاختبار (الخطية، والمئينية).

2. حساب الخطأ المعياري (Standard Error) للتحقق من دقة وفاعلية المعادلة الخطية الأفقية، للصور الجديدة .

3. حساب الخطأ المعياري(Standard Error) للتحقق من دقة وفاعلية المعادلة المئينية الأفقية، للصور الجديدة.

4. حساب معامل الصدق التقاطعي (Cross Validation) للتحقق من دقة وفاعلية المعادلة الخطية الأفقية، للصور الجديدة.

5. حساب معامل الصدق التقاطعي (Cross Validation) للتحقق من دقة وفاعلية المعادلة المئينية الأفقية، للصور الجديدة.

الفصل الرابع

نتائج الدراسة

سيتم في هذا الفصل وصف النتائج التي توصلت إليها الدراسة، والتي هدفت إلى معادلة صور مختارة من الامتحان الوطني للغة الإنجليزية في الأردن وفق نظرية الاختبار و نظرية الفقرة، ودراسة الخصائص السيكومترية لهذا الاختبار، وذلك من خلال الإجابة عن الأسئلة التالية:

السؤال الأول: ما هي الخصائص السيكومترية للصور الثلاث للامتحان الوطني للغة الإنجليزية في الأردن؟

السؤال الثاني: ما قيم القدرة المتعادلة في الصور الثلاث للامتحان الوطني للغة الإنجليزية في الأردن باستخدام نماذج نظرية الفقرة (أحادي المعلمة، وثنائي المعلمة، وثلاثي المعلمة)؟

السؤال الثالث: أي طرق المعادلة تحقق أكثر فاعلية؟

النتائج المتعلقة بالسؤال الأول:

فيما يتعلق بالسؤال الأول والمتعلق بدراسة الخصائص السيكومترية لصور الاختبار فقد تم حساب الإحصائيات الوصفية لدورات الامتحان و تحليل التباين الأحادي للدرجات الملاحظة، حيث تم حساب الإحصائيات الوصفية لصور الاختبار الثلاث، ويحتوي الجدول (5) عرضاً تفصيلياً لملخص الإحصائيات الوصفية لصور الاختبار الثلاث.

الجدول (5)

ملخص الإحصائيات الوصفية لصور الاختبار الثلاث

الإحصائي	الصورة		
	الأولى	الثانية	الثالثة
المتوسط الحسابي	42.06	41.27	41.44
الوسيط	40	39	39
المنوال	30	30	29
الانحراف المعياري	15.56	14.63	15.70
التباين	242.08	214.13	246.59
الالتواء	0.55	0.53	0.59
التفلطح	-0.34	-0.53	-0.37
المدى	77	70	74
العلامة الصغرى	13	14	13
العلامة العظمى	90	84	87
حجم العينة	979	854	899

يلاحظ من الجدول (5) تقارب قيم الإحصاءات الوصفية لصور الاختبار الثلاث.

تم استخدام اختبار ليفين (F) (Levine's) للتحقق فيما إذا كانت هناك فروق ذات دلالة إحصائية عند مستوى الدلالة (α=0.05) بين تباينات صور الاختبار الثلاث، فكانت قيمة (F) المحسوبة (2.291) و الدلالة الإحصائية لها (0.101). مما يعني عدم وجود فروق ذات دلالة إحصائية بين تباينات درجات المتقدمين تُعزى لصورة الاختبار. وهذا يعني تحقق شرط تجانس التباين.

وتم إجراء تحليل التباين الأحادي للبيانات للكشف عما إذا كانت هناك فروق ذات دلالة إحصائية عند مستوى الدلالة (α =0.05) بين المتوسطات الحسابية لدرجات المتقدمين تُعزى لصورة الاختبار، والجدول (6) يبين نتائج تحليل التباين الأحادي.

جدول (6)
تحليل التباين الأحادي لدرجات صور الاختبار الثلاث

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط مجموع المربعات	قيمة ف المحسوبة	الدلالة الإحصائية
بين الصور	321.172	2	160.586	0.684	0.505
داخل الصور	640845.104	2729	234.828		
الكلي	641166.276	2731			

يتضح من الجدول (6) عدم وجود فروق ذات دلالة إحصائية عند مستوى الدلالة (α=0.05) بين المتوسطات الحسابية لدرجات المتقدمين تُعزى لصورة الاختبار.

كما تم حساب مؤشرات ثبات صور الاختبار، حيث تم إيجاد قيم إحصائيات فقرات الاختبار حسب نظرية الاختبار من خلال تحليل البيانات لكل صورة من صور الاختبار، وحساب مؤشرات ثبات الاختبار وهي: معاملات ثبات الاتساق الداخلي لكل صورة من صور الاختبار الثلاث باستخدام معامل كرونباخ α (Cronbach α). وقد بلغت على التوالي 0.918، و0.906، و0.919، للصور الأولى والثانية والثالثة. ويلاحظ أن جميع معاملات الثبات وللصور الثلاث كانت مرتفعة.

كذلك تم حساب قيم مؤشر الخطأ المعياري للقياس للصور الثلاث وذلك من خلال المعادلة:

$$\sigma_e = (\sigma_x)\sqrt{1 - r_{xx'}}$$

وكانت قيمه على التوالي: 4.40، و4.39، و4.44 للصور الأولى والثانية والثالثة. ويلاحظ أن جميع قيم الخطأ المعياري للقياس متدنية نوعاً ما – أقل من 5%–.

كذلك تم حساب معاملات الصعوبة لكل فقرة وذلك بإيجاد نسبة المفحوصين الذين أجابوا عن الفقرة إجابة صحيحة من بين المفحوصين الذين حاولوا الإجابة عن هذه الفقرات، وتم حساب معاملات التمييز لكل فقرة (معامل الارتباط الثنائي بين الفقرة والاختبار ككل بعد حذف

الفقرة نفسها) باستخدام معامل الارتبـــاط الثنـــائي الخـــاص (Point Biserial Correlation) (r_{pbis})، والمتوسطات الحسابية لهذه المعاملات للنماذج الثلاث. و يبين الجـــدول (7) عرضــاً لمعاملات الصعوبة ومعاملات التمييز ودرجات فشر الزائية (Zr) المقابلـــة لمعـــاملات تمييـــز الفقرات والمتوسطات الحسابية والانحرافات المعيارية لمعاملات الصعوبة ولمعاملات التمييـــز ولصور الاختبار الثلاثة:

الجدول (7)

قيم معاملات الصعوبة و معاملات التمييز ودرجات فشر الزائية (Zr) المقابلة لمعاملات تمييز الفقرات في صور الاختبار الثلاث

الصورة	الأولى			الثانية			الثالثة		
رقم الفقرة	الصعوبة	التمييز	Z_r	الصعوبة	التمييز	Z_r	الصعوبة	التمييز	Z_r
1	0.281	0.333	0.346	0.487	0.482	0.532	0.317	0.124	0.328
2	0.343	0.224	0.357	0.226	0.300	0.230	0.238	0.225	0.243
3	0.336	0.221	0.349	0.151	0.069	0.152	0.323	0.234	0.335
4	0.463	0.371	0.501	0.171	0.095	0.173	0.339	0.267	0.353
5	0.343	0.322	0.357	0.320	0.356	0.331	0.318	0.203	0.330
6	0.263	0.211	0.270	0.316	0.245	0.327	0.357	0.174	0.374
7	0.482	0.431	0.525	0.329	0.380	0.342	0.197	0.096	0.199
8	0.245	0.310	0.250	0.356	0.274	0.372	0.331	0.290	0.344
9	0.464	0.424	0.503	0.354	0.431	0.370	0.523	0.418	0.580
10	0.291	0.034	0.299	0.663	0.492	0.798	0.456	0.380	0.492
11	0.293	0.362	0.302	0.363	0.474	0.380	0.314	0.251	0.325
12	0.449	0.262	0.483	0.330	0.268	0.343	0.314	0.287	0.325
13	0.308	0.369	0.319	0.245	0.308	0.250	0.469	0.506	0.509
14	0.153	0.124	0.154	0.252	0.026	0.257	0.301	0.304	0.311
15	0.534	0.302	0.595	0.457	0.511	0.493	0.154	0.094	0.155
16	0.342	0.363	0.356	0.429	0.263	0.458	0.222	0.249	0.226
17	0.432	0.275	0.462	0.349	0.138	0.364	0.251	-0.091	0.257
18	0.395	0.431	0.418	0.321	0.352	0.333	0.216	-0.106	0.219
19	0.308	0.422	0.319	0.672	0.423	0.815	0.229	0.141	0.233
20	0.528	0.421	0.587	0.326	0.420	0.338	0.263	0.269	0.269
21	0.779	0.304	1.042	0.639	0.303	0.757	0.712	0.448	0.891
22	0.444	0.493	0.477	0.434	0.378	0.465	0.464	0.477	0.502
23	0.532	0.477	0.592	0.344	0.401	0.359	0.457	0.255	0.494
24	0.651	0.480	0.777	0.457	0.229	0.493	0.253	0.058	0.258
25	0.345	0.128	0.360	0.301	0.414	0.311	0.127	0.094	0.127
26	0.271	0.234	0.278	0.189	0.223	0.191	0.519	0.446	0.576
27	0.353	0.312	0.369	0.396	0.402	0.419	0.314	0.091	0.325
28	0.459	0.473	0.496	0.465	0.402	0.504	0.386	0.411	0.407
29	0.335	0.292	0.348	0.444	0.174	0.477	0.650	0.471	0.775
30	0.320	0.133	0.332	0.270	0.194	0.277	0.390	0.365	0.412
31	0.571	0.362	0.650	0.536	-0.048	0.599	0.522	0.396	0.579
32	0.672	0.386	0.815	0.515	0.348	0.570	0.484	0.374	0.528
33	0.459	0.287	0.496	0.502	0.322	0.552	0.551	0.240	0.619
34	0.364	0.396	0.382	0.582	0.473	0.665	0.640	0.395	0.757

الثالثة			الثانية			الأولى			الصورة
Z_r	التمييز	الصعوبة	Z_r	التمييز	الصعوبة	Z_r	التمييز	الصعوبة	رقم الفقرة
0.446	0.312	0.418	0.606	0.532	0.541	0.188	0.165	0.186	35
0.521	0.519	0.478	0.621	0.423	0.552	0.576	0.263	0.519	36
0.436	0.211	0.410	0.394	0.339	0.375	0.734	0.410	0.626	37
0.341	0.452	0.328	0.715	0.226	0.614	0.447	0.328	0.419	38
0.344	0.391	0.331	0.749	0.489	0.635	0.311	0.399	0.301	39
0.138	0.242	0.137	0.523	0.366	0.480	0.276	0.292	0.269	40
0.353	0.418	0.339	0.808	0.482	0.669	0.189	0.232	0.187	41
0.550	0.505	0.501	0.619	0.432	0.550	0.149	0.054	0.148	42
0.275	0.152	0.268	0.859	0.285	0.696	0.225	0.157	0.221	43
0.227	0.251	0.224	0.187	0.066	0.185	0.245	0.153	0.240	44
0.506	0.433	0.467	0.421	0.383	0.398	0.402	0.296	0.382	45
0.698	0.496	0.603	0.589	0.160	0.529	0.722	0.450	0.618	46
0.712	0.419	0.612	0.416	0.133	0.393	0.293	0.319	0.285	47
0.759	0.428	0.641	0.331	0.125	0.320	0.261	0.216	0.255	48
0.437	0.291	0.412	0.325	0.307	0.314	0.229	0.157	0.226	49
0.346	0.348	0.333	0.334	0.206	0.322	0.810	0.361	0.669	50
0.632	0.366	0.560	0.464	0.440	0.433	0.478	0.299	0.445	51
0.366	0.326	0.350	0.523	0.388	0.480	0.481	0.414	0.447	52
0.508	0.251	0.468	0.498	0.280	0.460	0.522	0.487	0.480	53
0.631	0.431	0.558	0.229	0.256	0.225	0.879	0.462	0.706	54
0.189	0.099	0.187	0.594	0.283	0.533	0.729	0.428	0.622	55
0.386	0.249	0.368	0.705	0.458	0.608	0.227	0.140	0.223	56
0.336	0.251	0.324	0.830	0.495	0.680	0.573	0.437	0.517	57
0.837	0.350	0.684	0.621	0.383	0.552	0.467	0.122	0.436	58
0.460	0.245	0.430	0.599	0.278	0.536	0.948	0.243	0.739	59
0.476	0.500	0.443	0.683	0.528	0.594	0.576	0.292	0.519	60
0.360	0.180	0.345	0.327	0.233	0.316	0.704	0.392	0.607	61
0.451	0.399	0.423	0.516	0.312	0.474	0.534	0.460	0.489	62
0.252	0.306	0.247	0.823	0.441	0.677	0.354	0.287	0.340	63
0.347	0.364	0.334	0.414	0.452	0.392	0.623	0.414	0.553	64
0.535	0.300	0.489	0.787	0.41	0.657	0.670	0.318	0.585	65
0.555	0.404	0.504	0.438	0.455	0.412	0.246	-0.040	0.241	66
0.975	0.404	0.751	0.321	0.063	0.310	0.520	0.390	0.478	67
0.600	0.417	0.537	0.511	0.260	0.471	0.911	0.387	0.721	68
0.469	0.164	0.437	0.417	0.179	0.395	0.248	0.293	0.243	69
0.487	0.202	0.452	0.331	0.387	0.320	0.992	0.362	0.758	70
0.231	0.110	0.227	0.601	0.452	0.537	0.356	0.133	0.342	71
0.559	0.462	0.507	0.416	0.090	0.393	0.627	0.484	0.556	72
0.735	0.179	0.626	0.255	0.108	0.249	0.404	0.402	0.384	73
0.243	0.278	0.238	0.244	0.351	0.239	0.419	0.233	0.396	74
0.429	0.413	0.405	0.274	-0.076	0.267	0.504	0.258	0.465	75
0.477	0.592	0.444	0.220	0.042	0.217	0.846	0.288	0.689	76
0.588	0.297	0.528	0.229	-0.052	0.225	0.321	-0.073	0.310	77
0.511	0.429	0.471	0.307	0.096	0.297	0.377	0.385	0.360	78
0.715	0.356	0.614	0.119	-0.137	0.118	0.298	0.169	0.290	79
0.497	0.244	0.459	0.549	0.472	0.500	0.146	-0.077	0.145	80
1.065	0.366	0.788	0.284	0.057	0.276	0.704	0.295	0.607	81
0.443	0.218	0.416	0.601	0.361	0.537	0.353	0.219	0.339	82
0.379	0.370	0.362	0.643	0.218	0.567	0.553	0.343	0.503	83
0.327	0.280	0.316	0.440	0.304	0.413	0.260	0.308	0.254	84
0.299	0.346	0.290	0.270	0.201	0.263	0.363	0.346	0.348	85
0.288	0.232	0.280	0.546	0.164	0.498	0.584	0.225	0.526	86
0.473	0.259	0.440	0.293	0.048	0.285	0.569	0.328	0.514	87

الصورة	الأولى			الثانية			الثالثة		
رقم الفقرة	الصعوبة	التمييز	Z_r	الصعوبة	التمييز	Z_r	الصعوبة	التمييز	Z_r
88	0.637	0.409	0.753	0.288	0.041	0.296	0.115	0.005	0.115
89	0.592	0.493	0.680	0.158	0.039	0.159	0.370	0.115	0.389
90	0.254	0.238	0.260	0.715	0.395	0.898	0.526	0.334	0.585
91	0.282	0.242	0.289	0.383	0.328	0.403	0.251	0.258	0.257
92	0.453	0.278	0.489	0.645	0.231	0.767	0.478	0.330	0.521
93	0.208	0.185	0.211	0.371	0.167	0.390	0.372	0.255	0.390
94	0.380	0.318	0.400	0.315	0.340	0.326	0.614	0.433	0.715
95	0.313	-0.022	0.324	0.224	-0.093	0.227	0.586	0.311	0.672
96	0.357	0.332	0.374	0.563	0.389	0.638	0.434	0.366	0.465
97	0.490	0.491	0.536	0.276	0.090	0.284	0.791	0.308	1.074
98	0.496	0.436	0.544	0.292	0.177	0.300	0.468	0.381	0.508
99	0.655	0.302	0.784	0.560	0.344	0.632	0.514	0.432	0.568
100	0.365	0.244	0.383	0.313	0.065	0.323	0.517	0.358	0.573
المتوسط	**0.420**	**0.300**	**0.466**	**0.413**	**0.277**	**0.454**	**0.414**	**0.302**	**0.457**
الانحراف المعياري	**0.154**	**0.128**	**0.204**	**0.147**	**0.160**	**0.188**	**0.148**	**0.132**	**0.195**

وللكشف عما إذا كان الفرق بين متوسطات الصعوبة دال إحصائياً عند مستوى الدلالة (α=0.05) تُعزى لصورة الاختبار، فقد تم تحويل معاملات الصعوبة إلى قيم دلتا المناظرة والتي هي درجات معيارية محولة بوسط حسابي قدره (13) وانحراف معياري قدره (4)؛ وذلك لأن قيم دلتا المناظرة تجعل توزيع المعاينة قريباً من التوزيع الاعتدالي بغض النظر عن قيمة هذا المعامل في المجتمع أو حجم العينة، وحسب المعادلة التالية $\Delta = 4Z + 13$. وتم استخدام تحليل التباين الأحادي للكشف عن الفروق في متوسطات معاملات الصعوبة بين النماذج في الدورات الثلاث، و يبين الجدول (8) نتيجة التحليل.

الجدول(8)

تحليل التباين الأحادي لقيم دلتا المحولة عن قيم معاملات صعوبة الفقرات لصور الاختبار الثلاث

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط مجموع المربعات	قيمة F المحسوبة	الدلالة الإحصائية
بين المجموعات	.141	2	.070	.004	.996
داخل الفقرات	4745.525	297	15.978		
الكلي	4745.665	299			

يلاحظ من خلال الجدول (8) عدم وجود فروق ذات دلالة إحصائية بين متوسطات معاملات الصعوبة للفقرات في الصور الثلاث.

وللكشف عن مدى الفروق بين المتوسطات الحسابية لمعاملات التمييز لصور الاختبار، حيث تم تحويل معاملات التمييز للفقرات إلى درجات فشر الزائية (Z_r) لأن إجراء هذا التحويل الرياضي على معاملات التمييز يجعل توزيع المعاينة قريباً من التوزيع الاعتدالي بغض النظر عن قيمة هذا المعامل في المجتمع أو حجم العينة، ولمعرفة مدى الاختلاف بين متوسطات معاملات التمييز في صور الاختبار تم استخدام تحليل التباين الأحادي، و يبين الجدول (9) نتائج التحليل.

الجدول(9)

تحليل التباين الأحادي لقيم (Zr) المحولة عن قيم تمييز الفقرات في صور الاختبار الثلاث

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط مجموع المربعات	قيمة F المحسوبة	الدلالة الإحصائية
بين المجموعات	0.041	2	0.021	0.871	0.420
داخل الفقرات	7.031	285.716	0.024		
الكلي	7.072	299			

يلاحظ من خلال الجدول (9) عدم وجود فروق ذات دلالة إحصائية بين متوسطات معاملات التمييز للفقرات في الصور الثلاث.

النتائج المتعلقة بالسؤال الثاني:

وفيما يتعلق بالسؤال الثاني من أسئلة الدراسة والمتعلق بقيم القدرات المتعادلة في صور الاختبار الثلاث للامتحان الوطني للغة الإنجليزية في الأردن باستخدام نماذج نظرية الفقرة (أحادي المعلمة، وثنائي المعلمة، وثلاثي المعلمة)، فقد تم إعادة تدريج درجات القدرة للمتقدمين على الصور الثلاث للاختبار على مقياس مشترك، وذلك للحصول على قيم القدرة المتعادلة لصور الاختبار.

وتفترض نماذج نظرية الفقرة: وجود قدرة واحدة تفسر أداء المفحوص في الاختبار. وهذا الافتراض يسمى افتراض أحادية البعد (Unidimensionality). وفي هذه الدراسة استخدم الباحث أسلوب التحليل العاملي (Factor Analysis) للتحقق من افتراض أحادية البُعد.

فقد استخدم الباحث أسلوب تحليل المكونات الأساسية (Principal Component Analysis) لاستجابات المفحوصين على كل صورة من صور الاختبار الثلاث. وقد تم حساب قيمة الجذر الكامن (Eigen Value) ونسبة التباين المفسر (Explained Variance) لكل عامل من العوامل، وقد أجري التدوير المتعامد (Varimax Rotation) للعوامل التي كانت قيمة الجذر الكامن لها أكبر من واحد. و يبين الجدول (10) الجذور الكامنة ونسبة التباين المفسر لصور الاختبار الثلاث.

الجدول (10)
قيم الجذور الكامنة ونسب التباين المفسر المقابلة لعوامل صور الاختبار الثلاث

رقم العامل	الصورة الأولى		الصورة الثانية		للصورة الثالثة	
	الجذر الكامن	نسبة التباين المفسر	الجذر الكامن	نسبة التباين المفسر	الجذر الكامن	نسبة التباين المفسر
1	**12.468**	**12.468**	**12.113**	**12.113**	**12.687**	**12.687**
2	**2.778**	**2.778**	**2.957**	**2.957**	**2.810**	**2.810**
3	**2.366**	**2.366**	**2.095**	**2.095**	**2.593**	**2.593**
4	**2.022**	**2.022**	1.843	1.843	**2.093**	**2.093**
5	1.767	1.767	1.713	1.713	1.589	1.589
6	1.575	1.575	1.582	1.582	1.510	1.510
7	1.481	1.481	1.517	1.517	1.497	1.497
8	1.447	1.447	1.482	1.482	1.426	1.426
9	1.382	1.382	1.472	1.472	1.392	1.392
10	1.371	1.371	1.426	1.426	1.375	1.375
11	1.363	1.363	1.385	1.385	1.362	1.362
12	1.332	1.332	1.378	1.378	1.331	1.331
13	1.322	1.322	1.343	1.343	1.328	1.328
14	1.301	1.301	1.325	1.325	1.277	1.277
15	1.277	1.277	1.292	1.292	1.264	1.264
16	1.269	1.269	1.267	1.267	1.253	1.253
17	1.219	1.219	1.253	1.253	1.223	1.223
18	1.206	1.206	1.237	1.237	1.219	1.219

1.214	1.214	1.219	1.219	1.198	1.198	19
1.195	1.195	1.199	1.199	1.159	1.159	20
1.178	1.178	1.181	1.181	1.146	1.146	21
1.159	1.159	1.164	1.164	1.138	1.138	22
1.130	1.130	1.146	1.146	1.131	1.131	23
1.123	1.123	1.144	1.144	1.123	1.123	24
1.087	1.087	1.122	1.122	1.106	1.106	25
1.077	1.077	1.117	1.117	1.091	1.091	26
1.069	1.069	1.105	1.105	1.075	1.075	27
1.051	1.051	1.092	1.092	1.064	1.064	28
1.042	1.042	1.059	1.059	1.057	1.057	29
1.030	1.030	1.057	1.057	1.047	1.047	30
1.021	1.021	1.038	1.038	1.025	1.025	31
1.010	1.010	1.027	1.027	1.020	1.020	32
1.001	1.001	0.998	0.998	1.012	1.012	33

يتضح من الجدول (10) أن نسبة التباين المفسر لكل عامل كانت تساوي قيمة الجذر الكامن له وذلك لأن نسبة التباين المفسر = (قيمة التباين العائد للعامل/عدد العوامل) × 100% وحيث أن قيمة التباين العائد للمتغير = قيمة الجذر الكامن له، وعدد العوامل هنا = عدد الفقرات = 100 فقرة، فإن نسبة التباين المفسر للعامل = (قيمة الجذر الكامن/ 100)× 100% = قيمة الجذر الكامن.

كما يتضح من الجدول (10) أن أحادية البعد قد تحققت في صور الاختبار الثلاث؛ وذلك اعتماداً على محك الجذر الكامن كمؤشر على أحادية البعد، فكما هو واضح من الجدول (12) بالرغم من أن نسبة التباين المفسر للعامل الأول ليست مرتفعة – أقل من 20–، إلا أن الجذر الكامن للعامل الأول يطغى بشكل جليّ على الجذور الكامنة لبقية العوامل، حيث أن نسبة الجذر الكامن للعامل الأول إلى الجذر الكامن للعامل الثاني تزيد عن (2). حيث ذكر لورد (Lord) (المشار إليه في دراسة (Albanese and Forsyth, 1984)) بأن الفقرات تكون أحادية البعد إذا كانت قيمة الجذر الكامن للعامل الأول كبيرة بالمقارنة مع قيمة الجذر الكامن الثاني، بحيث تكون قيمة الجذر الكامن الثاني ليست أكبر بكثير من أي عامل آخر، إذ أن نسبة الجذر الكامن

للعامل الأول إلى الجذر الكامن الثاني كبيرة وتزيد عن (2). كما ويلاحظ أن نسبة التباين المفسر لكل من العوامل المتبقية متقاربة جداً، أي أن هناك شبه استقرار في نسب التباين المفسرة للعوامل جميعها باستثناء العامل الأول. وهذا يعني تحقق أحادية البعد في بيانات هذا الاختبار. لذا فإن هذا يُعد مؤشراً على أحادية بُعد الاختبار.

الاستقلال الموضعي Local item independence:

ينص هذا الافتراض على أن استجابات المفحوصين الذين لديهم نفس القدرة لفقرة ما مستقلة إحصائياً، أي أن استجابة المفحوص على فقرة ما يجب أن لا تؤثر في استجابته على أي فقرة أخرى، بمعنى أن إجابة الفقرة لا تُعطي تلميحات أو معيقات لإجابة فقرة أخرى. وهذا الافتراض يعني أن لا شيء يؤثر في التحصيل في الاختبار سوى قدرة المفحوص وخصائص الفقرة. وعندما يتحقق هذا الافتراض، يصبح احتمال الحصول على علامة لمجموعة من الفقرات هو حاصل ضرب الاحتمال لكل من هذه الفقرات. ويشير هامبلتون وسوامنيثان (Hambleton & Swaminatan, 1985) إلى أن هذا الافتراض يكافئ افتراض أحادية البعد، وهذا يعني أنه إذا تحقق افتراض أحادية البعد في الاختبار فإن الاختبار يحقق افتراض الاستقلال لموضعي. وهذا ما حصل في الدراسة الحالية.

افتراض تساوي القدرة التمييزية والتي يفترضها نموذج راش (النموذج أحادي المعلمة):

فكما أشار هامبلتون وسوامنيثان (Hambleton & Swaminatan, 1985) بأن هذا الافتراض يتحقق عندما تكون قيم مؤشرات التمييز واقعة ضمن الفترة (متوسط التمييز− 0.15، متوسط التمييز+0.15). وقد تم حساب معاملات التمييز لفقرات الاختبار من خلال

حساب معاملات الارتباط بين الأداء على الفقرة والدرجة الكلية لجميع فقرات الاختبار في كل صورة من صور الاختبار الثلاث كما هو مبين في الجدول (7).

وحيث كان متوسط معاملات التمييز (0.30) للصورة الأولى فإن حدود المطابقة للفقرات المطابقة تتراوح ما بين (0.15 و 0.45) وبالتالي فإن قيم معاملات الارتباط الثنائية الخاصة جميعها ضمن هذا المدى باستثناء (22) فقرة، والتي يتوقع أن يتم حذفها عند إجراء مطابقة الفقرات للنموذج. بينما كان متوسط معاملات التمييز للصورة الثانية (0.28)، وبالتالي فإن حدود المطابقة للفقرات المطابقة تتراوح ما بين (0.13 و 0.43) وبالتالي فإن قيم معاملات الارتباط الثنائية الخاصة جميعها ضمن هذا المدى باستثناء (37) فقرة، والتي يتوقع أن يتم حذفها عند إجراء مطابقة الفقرات للنموذج. كما كان متوسط معاملات التمييز للصورة الثالثة (0.30)، وبالتالي فإن حدود المطابقة للفقرات المطابقة تتراوح ما بين (0.15 و 0.45) وبالتالي فإن قيم معاملات الارتباط الثنائية الخاصة جميعها ضمن هذا المدى باستثناء (22) فقرة، والتي يتوقع أن يتم حذفها عند إجراء مطابقة الفقرات للنموذج. وكذلك يلاحظ أن قيمة الانحراف المعياري لهذه المعاملات صغيرة حيث كان الانحراف المعياري لمعاملات التمييز للدورة الأولى (0.13) وللدورة الثانية (0.16) وللدورة الثالثة (0.13)، ويُعد ذلك مؤشراً على تحقق هذا الافتراض.

أضف إلى ذلك أن تحليل التباين الأحادي بين عدم وجود فروق ذات دلالة إحصائية بين معاملات تمييز فقرات صور الاختبار الثلاث.

ومن الجدير بالذكر أن افتراضات نظرية الفقرة لن تتحقق بشكل مطلق في أي مجموعة من البيانات (Albanese & Foresyth, 1984).

وبعد ذلك تمت مطابقة الاستجابات عن فقرات كل صورة من صور الاختبار الثلاث لنماذج نظرية الفقرة الثلاث حيث تم تحليل استجابات أفراد عينة الدراسة الذين أجابوا عن فقرات كل صورة من صور الاختبار الثلاث باستخدام برنامج (BILOG) المصمم لتحليل البيانات وفقاً لنماذج نظرية الفقرة. و يبين الجدول (11) عدد الأفراد غير المطابقين لنماذج نظرية الفقرة وعدد الفقرات غير المطابقة لنماذج نظرية الفقرة ولصور الاختبار الثلاث.

الجدول (11)

عدد الأفراد غير المطابقين وعدد الفقرات غير المطابقة لنماذج نظرية الفقرة ولصور الاختبار الثلاث

النموذج	حالة المطابقة للأفراد والفقرات	الصورة		
		الأولى	الثانية	الثالثة
أحادي المعلمة	عدد الأفراد غير المطابقين	41	43	32
	عدد الفقرات غير المطابقة	58	65	55
ثنائي المعلمة	عدد الأفراد غير المطابقين	33	36	30
	عدد الفقرات غير المطابقة	13	14	21
ثلاثي المعلمة	عدد الأفراد غير المطابقين	35	36	32
	عدد الفقرات غير المطابقة	2	6	5

ويعد الأفراد غير مطابقين للنموذج؛ لأن استجاباتهم الملاحظة تكون بعيدة عن توقعات النموذج، كأن يجيب الفرد عن فقرة ما إجابة خاطئة على الرغم من أن صعوبة الفقرة دون مستوى قدرته، أو أن يجيب الفرد عن فقرة ما إجابة صحيحة على الرغم من أن مستوى صعوبة الفقرة أعلى من قدرته. أما الفقرات غير المطابقة للنموذج فهي الفقرات التي تبتعد الاستجابات الملاحظة عنها عن تلك التي يتوقعها النموذج وبشكل جوهري، مما يستدعي إعادة النظر بها لمعرفة أسباب عدم مطابقتها، فقد تكون هناك أسباب معينة مثل الصياغة أو الوضوح أو بدائل الإجابة . . . إلى غير ذلك من أسباب (الشريفين، 2009).

ثم تم التحقق من وجود اختلاف ذو دلالة إحصائية (α=0.05) بين نسب الفقرات المطابقة، وكذلك نسب الأفراد المطابقين لكل نموذج من نماذج نظرية الفقرة على مستوى صور الاختبار الثلاث باستخدام مؤشر كاي تربيع. و يبين الجدول (12) نتيجة الاختبار.

الجدول (12)

اختبار اختلاف نسبة الفقرات المطابقة والأفراد المطابقين لنماذج نظرية الفقرة

	قيمة مربع كاي	درجة الحرية	الدلالة الإحصائية
الفقرات	0.356	2	0.837
الأفراد	6.065	2	0.048

يتضح من الجدول (12) عدم وجود اختلاف ذو دلالة إحصائية عند مستوى الدلالة (α=0.05) بين نسب الفقرات المطابقة لنماذج نظرية الفقرة على مستوى صور الاختبار الثلاث.

بينما يتضح من الجدول (12) وجود اختلاف ذو دلالة إحصائية عند مستوى الدلالة (α=0.05) بين نسب الأفراد المطابقين لنماذج نظرية الفقرة على مستوى صور الاختبار الثلاث. مما يؤيد فكرة معادلة نماذج أو صور الاختبار.

وأخيراً تمت معادلة نماذج الاختبار وفقاً لنماذج نظرية الفقرة، حيث تم تقدير قدرات الأفراد على كل صورة من صور الاختبار الثلاث وفق نماذج نظرية الفقرة، وذلك بعد حذف الأفراد والفقرات غير المطابقة لنماذج الفقرة. و يبين الجدول (13) تقدير القدرة بوحدة اللوجيت للأفراد، وقيم الأخطاء المعيارية في قياسها، وفقاً للنموذج أحادي المعلمة، ولكل صورة من صور الاختبار والمقابلة للدرجة الخام نفسها.

الجدول (13)

تقديرات قدرات الأفراد بوحدة اللوجيت وقيم الأخطاء المعيارية في قياسها لكل صورة من صور الاختبار الثلاث وفق النموذج أحادي المعلمة

الدرجة الخام	1pl					
	الصورة الأولى		الصورة الثانية		الصورة الثالثة	
	القدرة	الخطأ المعياري للقدرة	القدرة	الخطأ المعياري للقدرة	القدرة	الخطأ المعياري المعادل
2	-3.213	0.737				
3	-2.765	0.613	-2.530	0.556		
4	-2.435	0.540	-2.199	0.490	-2.445	0.531
5	-2.170	0.492	-1.930	0.446	-2.189	0.483
6	-1.946	0.457	-1.702	0.414	-1.974	0.449
7	-1.750	0.430	-1.501	0.390	-1.787	0.423
8	-1.575	0.409	-1.319	0.371	-1.620	0.402
9	-1.414	0.392	-1.152	0.356	-1.468	0.386
10	-1.266	0.379	-0.996	0.343	-1.327	0.372
11	-1.127	0.367	-0.849	0.333	-1.196	0.361
12	-0.996	0.358	-0.708	0.324	-1.073	0.352
13	-0.871	0.350	-0.573	0.317	-0.955	0.344
14	-0.751	0.343	-0.441	0.311	-0.843	0.338
15	-0.635	0.338	-0.312	0.306	-0.735	0.332
16	-0.522	0.333	-0.186	0.302	-0.630	0.328
17	-0.412	0.330	-0.060	0.299	-0.528	0.324
18	-0.304	0.327	0.065	0.297	-0.428	0.322
19	-0.198	0.325	0.190	0.295	-0.330	0.320
20	-0.092	0.324	0.317	0.294	-0.234	0.319
21	0.012	0.323	0.445	0.293	-0.138	0.318
22	0.117	0.324	0.577	0.293	-0.043	0.318
23	0.222	0.324	0.712	0.294	0.052	0.319
24	0.328	0.326	0.852	0.296	0.147	0.320
25	0.435	0.328	0.999	0.298	0.242	0.323
26	0.543	0.331	1.154	0.300	0.338	0.326
27	0.654	0.335	1.320	0.304	0.436	0.330
28	0.768	0.340	1.501	0.309	0.535	0.335
29	0.886	0.346	1.701	0.314	0.637	0.341
30	1.009	0.354	1.928	0.321	0.741	0.348
31	1.137	0.363	2.195	0.329	0.848	0.357
32	1.273	0.374			0.960	0.367
33	1.417	0.387			1.076	0.381
34	1.573	0.403			1.198	0.397
35	1.744	0.424			1.328	0.417
36	1.935	0.450			1.467	0.443
37	2.153	0.485			1.618	0.477
38	2.411	0.534			1.784	0.525
39	2.733	0.606			1.969	0.596
40	3.172	0.730				
41	3.898	1.016				
42					2.753	0.999
المتوسط الحسابي	**0.099**	**0.424**	**-0.086**	**0.338**	**-0.049**	**0.395**
الانحراف المعياري	**2.686**	**0.145**	**1.275**	**0.065**	**1.260**	**0.124**

ويتضح من الجدول (13) أن معالم القدرة الفعلية للأفراد قد تراوحت بين (-3.213) لوجيت و (3.898) لوجيت وبمتوسط حسابي (0.099) وبخطأ معياري في قياسها قدره (0.737، 1.016) على الترتيب للصورة الأولى للاختبار. كما تراوحت بين (-2.530) لوجيت و (2.195) لوجيت وبمتوسط حسابي (-0.086) وبخطأ معياري في قياسها قدره (0.556

0.329، على الترتيب للصورة الثانية للاختبار. و تراوحت بين (-2.445) لوجيت و (2.753) لوجيت وبمتوسط حسابي (-0.049) وبخطأ معياري في قياسها قدره (0.531 ،0.999) علـــى الترتيب للصورة الثالثة للاختبار.

و يبين الجدول (14) تقدير القدرة بوحدة اللوجيت للأفراد، وقيم الأخطاء المعيارية فـــي قياسها، وفقاً للنموذج ثنائي المعلمة، ولكل صورة من صور الاختبار والمقابلة للدرجـــة الخـــام نفسها.

الجدول (14)

تقديرات قدرات الأفراد بوحدة اللوجيت وقيم الأخطاء المعيارية في قياسها لكل صورة من صور الاختبار الثلاث وفق النموذج ثنائي المعلمة

العلامة الخام	2pl					
	الصورة الأولى		الصورة الثانية		الصورة الثالثة	
	القدرة	الخطأ المعياري للقدرة	القدرة	الخطأ المعياري للقدرة	القدرة	الخطأ المعياري للقدرة
8	-2.136	0.414				
9					-2.002	0.388
10	-1.940	0.392			-2.237	0.390
11	-1.946	0.387			-1.650	0.392
12			-1.943	0.379	-1.670	0.385
13	-1.793	0.367	-1.722	0.363	-1.483	0.371
14	-1.604	0.358	-1.726	0.341	-1.486	0.349
15	-1.555	0.340	-1.675	0.337	-1.502	0.344
16	-1.479	0.332	-1.520	0.328	-1.301	0.335
17	-1.396	0.324	-1.467	0.321	-1.211	0.328
18	-1.293	0.314	-1.346	0.311	-1.096	0.318
19	-1.246	0.310	-1.306	0.307	-1.006	0.314
20	-1.171	0.304	-1.194	0.301	-0.946	0.308
21	-1.082	0.298	-1.175	0.295	-0.901	0.301
22	-0.981	0.291	-1.050	0.288	-0.808	0.295
23	-0.914	0.287	-0.973	0.284	-0.703	0.291
24	-0.882	0.285	-0.878	0.282	-0.631	0.289
25	-0.803	0.281	-0.829	0.279	-0.585	0.285
26	-0.685	0.276	-0.750	0.275	-0.522	0.280
27	-0.660	0.275	-0.644	0.271	-0.437	0.279
28	-0.522	0.271	-0.564	0.269	-0.391	0.274
29	-0.478	0.270	-0.525	0.267	-0.283	0.273
30	-0.395	0.268	-0.418	0.265	-0.214	0.271
31	-0.329	0.267	-0.404	0.264	-0.156	0.270
32	-0.253	0.266	-0.266	0.263	-0.090	0.269
33	-0.161	0.265	-0.192	0.262	0.004	0.268
34	-0.096	0.265	-0.121	0.262	0.086	0.268
35	-0.036	0.265	-0.024	0.262	0.161	0.268
36	0.036	0.266	0.012	0.263	0.209	0.269
37	0.068	0.266	0.084	0.263	0.313	0.269

العلامة الخام	2pl					
	الصورة الأولى		الصورة الثانية		الصورة الثالثة	
	القدرة	الخطأ المعياري للقدرة	القدرة	الخطأ المعياري للقدرة	القدرة	الخطأ المعياري للقدرة
38	0.183	0.268	0.226	0.265	0.324	0.271
39	0.238	0.269	0.178	0.266	0.368	0.272
40	0.310	0.270	0.259	0.267	0.509	0.273
41	0.398	0.273	0.387	0.270	0.590	0.276
42	0.449	0.274	0.401	0.271	0.649	0.277
43	0.504	0.276	0.495	0.273	0.700	0.279
44	0.560	0.278	0.589	0.275	0.779	0.281
45	0.638	0.281	0.660	0.278	0.841	0.284
46	0.722	0.284	0.811	0.281	0.907	0.288
47	0.764	0.286	0.783	0.283	1.069	0.290
48	0.913	0.294	0.873	0.291	1.124	0.297
49	0.940	0.295	1.000	0.292	1.217	0.298
50	1.041	0.300	1.040	0.297	1.216	0.304
51	1.076	0.303	1.181	0.299	1.322	0.306
52	1.143	0.307	1.214	0.303	1.420	0.310
53	1.246	0.313	1.222	0.310	1.464	0.317
54	1.337	0.319	1.332	0.316	1.631	0.323
55	1.362	0.321	1.428	0.317	1.651	0.324
56	1.487	0.330	1.613	0.326	1.830	0.333
57	1.545	0.334	1.557	0.330	1.845	0.338
58	1.646	0.341	1.747	0.338	1.979	0.345
59	1.730	0.348	1.919	0.345	2.098	0.352
60	1.873	0.360	1.968	0.356	2.332	0.364
61	1.988	0.370	2.020	0.366	2.487	0.374
62	2.038	0.374	2.184	0.370	2.447	0.379
63	2.090	0.379	2.453	0.375	2.546	0.383
64	2.220	0.392	2.235	0.388	2.781	0.396
65	2.215	0.390			3.082	0.395
66	2.448	0.414	2.656	0.410	3.338	0.419
67	2.703	0.441	2.937	0.436	3.289	0.446
68	2.637	0.434	2.977	0.429	3.254	0.439
69	2.613	0.431	3.183	0.426		
70	2.968	0.471	3.427	0.466		
71	3.129	0.492				
72	3.218	0.502				
73	3.524	0.543				
74	3.130	0.491				
75	3.452	0.533				
76	3.806	0.583				
المتوسط الحسابي	0.605	0.338	0.420	0.312	0.476	0.320
الانحراف المعياري	1.579	0.079	1.416	0.052	1.467	0.049

ويتضح من الجدول (14) أن معالم القدرة الفعلية للأفراد قد تراوحت بين (-3.213) لوجيت و (3.806) لوجيت وبمتوسط حسابي (0.605) وبخطأ معياري في قياسها قدره (1.579، 0.583) على الترتيب للصورة الأولى للاختبار. كما تراوحت بين (-1.943) لوجيت و (3.427) لوجيت وبمتوسط حسابي (0.240) وبخطأ معياري في قياسها قدره (0.379، 0.466)

على الترتيب للصورة الثانية للاختبار. و تراوحت بين (2.002-) لوجيت و (3.254) لوجيت وبمتوسط حسابي (0.476) وبخطأ معياري في قياسها قدره (0.388، 0.439) على الترتيب للصورة الثالثة للاختبار.

والجدول يبين (15) تقدير القدرة بوحدة اللوجيت للأفراد، وقيم الأخطاء المعيارية في قياسها، وفقاً للنموذج ثلاثي المعلمة، ولكل صورة من صور الاختبار والمقابلة للدرجة الخام نفسها.

الجدول (15)

تقديرات قدرات الأفراد بوحدة اللوجيت وقيم الأخطاء المعيارية في قياسها لكل صورة من صور الاختبار الثلاث وفق النموذج ثلاثي المعلمة

العلامة الخام	3Pl الصورة الأولى القدرة	الصورة الأولى الخطأ المعياري للقدرة	الصورة الثانية القدرة المعادلة	الصورة الثانية الخطأ المعياري للقدرة	الصورة الثالثة القدرة المعادلة	الصورة الثالثة الخطأ المعياري للقدرة
11	-2.522	2.215				
13	-2.484	2.146	-1.904	1.478	-1.615	2.297
14	-2.508	2.189			-1.619	2.342
15	-2.492	2.160	-1.880	1.487		
16	-2.485	2.148	-1.907	1.479	-1.592	2.298
17	-2.454	2.092	-1.867	1.441	-1.600	2.238
18	-2.670	2.813	-1.867	1.937	-1.585	3.010
19	-2.716	2.846	-1.854	1.960	-1.586	3.045
20	-2.772	3.163	-1.905	2.178	-1.582	3.384
21	-1.630	1.197	-1.959	0.824	-1.738	1.280
22	-1.406	0.922	-1.362	0.635	-1.525	0.987
23	-1.417	0.921	-1.179	0.634	-1.207	0.985
24	-1.156	0.755	-0.936	0.520	-1.171	0.808
25	-1.245	0.892	-1.009	0.614	-1.027	0.954
26	-0.974	0.633	-0.790	0.436	-1.023	0.677
27	-1.078	0.681	-0.736	0.469	-0.715	0.729
28	-0.786	0.523	-0.392	0.360	-0.625	0.560
29	-0.564	0.402	-0.485	0.277	-0.577	0.430
30	-0.662	0.450	-0.259	0.310	-0.377	0.482
31	-0.284	0.299	-0.109	0.206	-0.198	0.320
32	-0.215	0.276	-0.125	0.190	-0.242	0.295
33	-0.159	0.264	0.045	0.182	-0.335	0.283
34	-0.129	0.263	0.053	0.181	-0.023	0.281
35	-0.037	0.237	0.146	0.163	0.031	0.254
36	0.072	0.218	0.270	0.150	0.139	0.233
37	0.092	0.213	0.299	0.147	0.126	0.228
38	0.180	0.200	0.390	0.138	0.261	0.214
39	0.234	0.193	0.351	0.133	0.293	0.206
40	0.244	0.192	0.425	0.132	0.373	0.205
41	0.323	0.181	0.474	0.125	0.425	0.194
42	0.389	0.174	0.501	0.120	0.475	0.186
43	0.401	0.172	0.513	0.119	0.498	0.184
44	0.417	0.171	0.556	0.118	0.530	0.183
45	0.469	0.165	0.598	0.114	0.585	0.177
46	0.572	0.156	0.638	0.107	0.573	0.167
47	0.576	0.156	0.634	0.107	0.599	0.167
48	0.628	0.151	0.721	0.104	0.654	0.162
49	0.613	0.153	0.768	0.105	0.723	0.163

3Pl						
الصورة الثالثة		الصورة الثانية		الصورة الأولى		العلامة الخام
الخطأ المعياري للقدرة	القدرة المعادلة	الخطأ المعياري للقدرة	القدرة المعادلة	الخطأ المعياري للقدرة	القدرة	
0.157	0.712	0.101	0.812	0.147	0.688	50
0.154	0.795	0.099	0.835	0.144	0.726	51
0.152	0.772	0.098	0.828	0.142	0.749	52
0.150	0.824	0.096	0.880	0.140	0.784	53
0.147	0.825	0.095	0.932	0.138	0.817	54
0.145	0.900	0.093	0.947	0.135	0.853	55
0.142	0.950	0.091	0.999	0.133	0.895	56
0.140	0.975	0.090	0.970	0.131	0.927	57
0.138	1.011	0.089	1.016	0.129	0.967	58
0.137	1.000	0.088	1.063	0.128	0.976	59
0.135	1.043	0.087	1.059	0.126	1.012	60
0.133	1.065	0.085	1.155	0.124	1.058	61
0.131	1.127	0.084	1.180	0.122	1.096	62
0.130	1.169	0.084	1.139	0.121	1.116	63
0.127	1.128	0.082	1.238	0.119	1.191	64
0.129	1.176	0.083	1.257	0.120	1.146	65
0.127	1.220	0.081	1.276	0.118	1.216	66
0.125	1.307	0.081	1.330	0.117	1.272	67
0.125	1.239	0.081	1.350	0.117	1.285	68
0.127	1.301	0.082	1.337	0.118	1.205	69
0.125	1.351	0.080	1.419	0.117	1.345	70
0.125	1.401	0.080	1.435	0.117	1.369	71
0.125	1.458	0.080	1.446	0.117	1.380	72
0.125	1.477	0.081	1.574	0.117	1.383	73
0.127	1.418	0.082	1.508	0.119	1.513	74
0.127	1.552	0.082	1.610	0.119	1.486	75
0.128	1.562	0.082	1.621	0.119	1.518	76
0.128	1.599			0.120	1.528	77
0.131	1.683	0.084	1.759	0.122	1.605	78
0.130	1.527	0.083	1.772	0.121	1.575	79
0.136	1.687	0.088	1.866	0.127	1.699	80
0.141	1.815	0.090	1.938	0.131	1.756	81
0.134	1.832			0.125	1.654	82
0.152	1.768			0.142	1.886	83
				0.146	1.925	84
				0.152	1.975	85
				0.166	2.081	87
				0.156	2.009	88
				0.198	2.262	89
0.503	0.386	0.334	0.364	0.495	0.290	المتوسط الحسابي
0.776	1.074	0.507	1.129	0.749	1.378	الانحراف المعياري

ويتضح من الجدول (15) أن معالم القدرة الفعلية للأفراد قد تراوحت بين (-2.522) لوجيت و (2.262) لوجيت وبمتوسط حسابي (0.290) وبخطأ معياري في قياسها قدره (2.215، 0.198) على الترتيب للصورة الأولى للاختبار. كما تراوحت بين (-1.904) لوجيت و (1.938) لوجيت وبمتوسط حسابي (0.364) وبخطأ معياري في قياسها قدره (1.478، 0.090) على الترتيب للصورة الثانية للاختبار. و تراوحت بين (-1.615) لوجيت و (1.768) لوجيت وبمتوسط حسابي (0.386) وبخطأ معياري في قياسها قدره (2.297، 0.152) على الترتيب للصورة الثالثة للاختبار.

يلاحظ من الجداول (13)، (14)، (15) أن بعض الخلايا فارغة، وهذه الخلايا هي الخلايا التي تقابل درجات خام لم يحصل عليها أي مفحوص من الذين تقدموا لهذه الصورة من الاختبار.

ثم تمت معادلة قيم تقديرات القدرة للأفراد في نماذج الصورتين الثانية والثالثة إلى قيم تقديرات القدرة للأفراد في صورة الدورة الأولى وفق النموذج أحادي المعلمة (نموذج راش) باستخدام المعادلة التالية المشار إليها في رايت وماسترز (Wright & Masters, 1982):

$$\theta_{eq} = \frac{\sigma_{eq}}{\sigma_A}\left(\theta_A - \overline{\theta_A}\right) + \bar{\theta}_{eq}$$

حيث:

θ_{eq}: تقديرات قدرات للأفراد في صورة الاختبار المعادل إلى تقدير قدرات الأفراد في نموذج الاختبار A.

σ_{eq}: الانحراف المعياري المعدل من الخطأ لتقديرات القدرة للأفراد في صورة الاختبار المعادل. ويتم حسابه من خلال المعادلة:

$$\sigma_{eq} = \sqrt{\sigma_A^2 - MSE}$$

حيث MSE_p: متوسط مربعات الأخطاء المعيارية في قياس قدرات الأفراد (Mean Square Measurement Error) ويحسب من العلاقة التالية:

$$MSE = \frac{\sum_{n=1}^{N} S_n^2}{N}$$

حيث:

S_n: الخطأ المعياري في تقدير قدرات الأفراد،

N: عدد الأفراد

σ_A: الانحراف المعياري المعدل من الخطأ لتقدير القدرة للأفراد في صورة الاختبار A.

θ_A: تقديرات قدرة الأفراد في صورة الاختبار A.

$\overline{\theta}_{eq}$: متوسط تقديرات قدرة الأفراد في صورة الاختبار المعادل.

$\overline{\theta}_A$: متوسط تقديرات قدرة الأفراد في صورة الاختبار A.

والجدول يبين (16) تقديرات القدرة بوحدة اللوجيت المتعادلة لصور الاختبار باستخدام النموذج الأحادي المعلمة (نموذج راش) والمقابلة للدرجة الخام الخاصة بنموذج الصورة الأولى.

جدول (16)

تقديرات القدرات المتعادلة لصور الاختبار الثلاث وفقاً للنموذج أحادي المعلمة

الدرجة الخام	1pl الصورة الأولى: القدرة	الصورة الثانية: القدرة المعادلة	الصورة الثالثة: القدرة المعادلة
2	-3.213	-2.881	-3.163
3	-2.765	-2.475	-2.723
4	-2.435	-2.176	-2.399
5	-2.170	-1.936	-2.138
6	-1.946	-1.733	-1.918
7	-1.750	-1.555	-1.725
8	-1.575	-1.395	-1.552
9	-1.414	-1.250	-1.395
10	-1.266	-1.116	-1.249
11	-1.127	-0.990	-1.112
12	-0.996	-0.871	-0.983
13	-0.871	-0.757	-0.860
14	-0.751	-0.648	-0.742
15	-0.635	-0.543	-0.628
16	-0.522	-0.441	-0.518
17	-0.412	-0.341	-0.410
18	-0.304	-0.243	-0.304
19	-0.198	-0.147	-0.199
20	-0.092	-0.051	-0.095
21	0.012	0.044	0.008
22	0.117	0.139	0.110
23	0.222	0.234	0.214
24	0.328	0.330	0.318
25	0.435	0.427	0.423
26	0.543	0.525	0.530
27	0.654	0.626	0.639
28	0.768	0.729	0.751
29	0.886	0.836	0.867

0.987	0.947	1.009	30
1.113	1.064	1.137	31
1.247	1.187	1.273	32
1.389	1.318	1.417	33
1.542	1.459	1.573	34
1.710	1.614	1.744	35
1.898	1.787	1.935	36
2.112	1.985	2.153	37
2.366	2.219	2.411	38
2.682	2.511	2.733	39
3.114	2.909	3.172	40
3.828	3.568	3.898	41

ثم تمت معادلة قيم تقديرات القدرة للأفراد في الصورتين الثانية والثالثة إلى قيم تقديرات القدرة للأفراد في الصورة الأولى وفق النموذج ثنائي المعلمة باستخدام نفس المعادلة المشار إليها سابقاً، و يبين الجدول (17) تقديرات القدرة بوحدة اللوجيت المتعادلة لصور الاختبار باستخدام النموذج الثنائي المعلمة والمقابلة للدرجة الخام الخاصة بالصورة الأولى.

الجدول (17)

تقديرات القدرات المتعادلة لصور الاختبار الثلاث وفقاً للنموذج ثنائي المعلمة

2pl			
الصورة الثالثة	الصورة الثانية	الصورة الأولى	الدرجة الخام
القدرة المعادلة	القدرة المعادلة	القدرة	
-2.155	-2.112	-2.136	8
-1.957	-1.918	-1.940	10
-1.963	-1.924	-1.946	11
-1.809	-1.773	-1.793	13
-1.617	-1.585	-1.604	14
-1.567	-1.537	-1.555	15
-1.490	-1.462	-1.479	16
-1.406	-1.379	-1.396	17
-1.302	-1.278	-1.293	18
-1.255	-1.232	-1.246	19
-1.180	-1.158	-1.171	20
-1.089	-1.069	-1.082	21
-0.987	-0.970	-0.981	22
-0.919	-0.903	-0.914	23
-0.887	-0.872	-0.882	24
-0.807	-0.793	-0.803	25
-0.687	-0.676	-0.685	26
-0.662	-0.651	-0.660	27
-0.523	-0.516	-0.522	28
-0.478	-0.472	-0.478	29
-0.395	-0.390	-0.395	30
-0.327	-0.324	-0.329	31
-0.251	-0.249	-0.253	32
-0.158	-0.158	-0.161	33
-0.092	-0.094	-0.096	34

2pl			
الصورة الثالثة	الصورة الثانية	الصورة الأولى	الدرجة الخام
القدرة المعادلة	القدرة المعادلة	القدرة	
-0.031	-0.035	-0.036	35
0.042	0.037	0.036	36
0.074	0.068	0.068	37
0.190	0.182	0.183	38
0.246	0.236	0.238	39
0.319	0.308	0.310	40
0.408	0.395	0.398	41
0.460	0.446	0.449	42
0.515	0.500	0.504	43
0.571	0.555	0.560	44
0.650	0.632	0.638	45
0.736	0.716	0.722	46
0.778	0.757	0.764	47
0.929	0.905	0.913	48
0.956	0.931	0.940	49
1.058	1.031	1.041	50
1.094	1.066	1.076	51
1.162	1.133	1.143	52
1.265	1.234	1.246	53
1.358	1.325	1.337	54
1.383	1.349	1.362	55
1.509	1.472	1.487	56
1.568	1.530	1.545	57
1.671	1.630	1.646	58
1.756	1.713	1.730	59
1.900	1.855	1.873	60
2.016	1.968	1.988	61
2.066	2.017	2.038	62
2.119	2.069	2.090	63
2.251	2.198	2.220	64
2.245	2.193	2.215	65
2.481	2.423	2.448	66
2.739	2.676	2.703	67
2.672	2.610	2.637	68
2.648	2.586	2.613	69
3.007	2.938	2.968	70
3.170	3.097	3.129	71
3.261	3.186	3.218	72
3.570	3.488	3.524	73
3.171	3.098	3.130	74
3.497	3.417	3.452	75
3.855	3.767	3.806	76

وأخيراً تمت معادلة قيم تقديرات القدرة للأفراد في صورتي الاختبار الثانية والثالثة إلى قيم تقديرات القدرة للأفراد في صورة الاختبار الأولى وفق النموذج ثلاثي المعلمة باستخدام نفس المعادلة المشار إليها سابقاً، و يبين الجدول (18) تقديرات القدرة بوحـــدة اللوجيـــت المتعادلـــة

لنماذج الاختبار باستخدام النموذج الثلاثي المعلمة والمقابلة للدرجة الخـــام الخاصــــة بـــصورة الاختبار الأولى.

الجدول (18)

تقديرات القدرات المتعادلة لصور الاختبار الثلاث وفقاً للنموذج ثلاثي المعلمة

3Pl			
الصورة الثالثة	الصورة الثانية	الصورة الأولى	الدرجة الخام
القدرة المعادلة	القدرة المعادلة	القدرة	
-2.648	-1.704	-2.522	11
-2.607	-1.678	-2.484	13
-2.633	-1.694	-2.508	14
-2.615	-1.683	-2.492	15
-2.609	-1.678	-2.485	16
-2.575	-1.657	-2.454	17
-2.806	-1.806	-2.670	18
-2.856	-1.837	-2.716	19
-2.915	-1.875	-2.772	20
-1.693	-1.089	-1.630	21
-1.453	-0.935	-1.406	22
-1.465	-0.943	-1.417	23
-1.186	-0.763	-1.156	24
-1.281	-0.824	-1.245	25
-0.992	-0.638	-0.974	26
-1.103	-0.709	-1.078	27
-0.790	-0.508	-0.786	28
-0.552	-0.355	-0.564	29
-0.657	-0.423	-0.662	30
-0.254	-0.163	-0.284	31
-0.179	-0.115	-0.215	32
-0.120	-0.077	-0.159	33
-0.087	-0.056	-0.129	34
0.011	0.008	-0.037	35
0.128	0.083	0.072	36
0.149	0.096	0.092	37
0.243	0.157	0.180	38
0.301	0.194	0.234	39
0.311	0.201	0.244	40
0.396	0.255	0.323	41
0.467	0.301	0.389	42
0.480	0.309	0.401	43
0.497	0.320	0.417	44
0.553	0.356	0.469	45
0.662	0.427	0.572	46
0.667	0.429	0.576	47
0.723	0.466	0.628	48
0.706	0.455	0.613	49
0.787	0.507	0.688	50
0.828	0.533	0.726	51
0.852	0.548	0.749	52
0.889	0.573	0.784	53
0.925	0.595	0.817	54
0.964	0.621	0.853	55
1.009	0.650	0.895	56
1.042	0.671	0.927	57
1.085	0.699	0.967	58
1.095	0.705	0.976	59
1.134	0.730	1.012	60

الدرجة الخام	3Pl الصورة الأولى		
	الصورة الأولى	الصورة الثانية	الصورة الثالثة
	القدرة	القدرة المعادلة	القدرة المعادلة
61	1.058	0.761	1.182
62	1.096	0.788	1.224
63	1.116	0.801	1.245
64	1.191	0.853	1.325
65	1.146	0.822	1.277
66	1.216	0.870	1.352
67	1.272	0.909	1.411
68	1.285	0.918	1.425
69	1.205	0.863	1.340
70	1.345	0.959	1.490
71	1.369	0.976	1.515
72	1.380	0.983	1.527
73	1.383	0.985	1.531
74	1.513	1.075	1.669
75	1.486	1.056	1.641
76	1.518	1.079	1.675
77	1.528	1.085	1.686
78	1.605	1.138	1.768
79	1.575	1.117	1.735
80	1.699	1.203	1.868
81	1.756	1.242	1.930
82	1.654	1.172	1.820
83	1.886	1.332	2.069
84	1.925	1.359	2.110
85	1.975	1.393	2.164
87	2.081	1.466	2.277
88	2.009	1.416	2.200
89	2.262	1.591	2.471

وتم استخدام تحليل التباين الأحادي للكشف عما إذا كانت هناك فروق ذات دلالة إحصائية بين متوسطات القدرة المتعادلة وفق كل نموذج من نماذج نظرية الفقرة الثلاث عند مستوى الدلالة ($\alpha=0.05$) و يبين الجدول (19) نتائج تحليل التباين الأحادي.

الجدول (19)

تحليل التباين الأحادي للقدرات المتعادلة لصور الاختبار الثلاث وفق نماذج نظرية الفقرة

النموذج	مصدر التباين	مجموع المربعات	درجة الحرية	متوسط مجموع المربعات	قيمة ف المحسوبة	الدلالة الإحصائية
1PL	بين الصور	3.417	2	1.709	2.544	0.079
	داخل الصور	1753.349	2611	0.672		
	الكلي	1756.767	2613			
2PL	بين الصور	0.013	2	0.006	0.006	0.994
	داخل الصور	2979.948	2604	1.144		
	الكلي	2979.961	2606			
3PL	بين الصور	1.207	2	0.604	0.577	0.561
	داخل الصور	2700.178	2583	1.045		
	الكلي	2701.385	2585			

يتضح من خلال الجدول (19) عدم وجود فروق ذات دلالة إحصائية بـين متوسـطات القدرات المقدرة عند مستوى الدلالة (α=0.05) وفق كل نموذج من نماذج نظرية الفقرة الثلاث.

وفيما يتعلق بالتحقق من مدى تحقق دقة معادلة الاختبارات وفاعليتها باسـتخدام محـك الصدق التقاطعي لدقة المعادلة (Cross-Validation Criterion) و الذي يبين درجة استقرار العلامات المكافئة أو المعادلة في عينتين عشوائيتين مستقلتين عن بعضهما الـبعض، فقـد تـم اختيار عينة عشوائية من صورة الاختبار الأولى وعينتين عشوائيتين مـستقلتين مـن صـورة الاختبار الثانية، وعينتين عشوائيتين مستقلتين من صورة الاختبار الثالثة كـل منهمـا تحـوي (200) فرد من مجتمع الدراسة، ثم تمت معادلة قدرات هذه العينات وفق نماذج نظري الفقـرة، ثم حساب معاملات الصدق التقاطعي وفق المعادلة السابقة. و يبين الجدول (20) قيم معـاملات الصدق التقاطعي:

الجدول (20)
معاملات الصدق التقاطعي للمعادلة وفق نماذج استجابة الفقرة

الصور المعادلة	النموذج		
	أحادي المعلمة	ثنائي المعلمة	ثلاثي المعلمة
	معامل الصدق التقاطعي	معامل الصدق التقاطعي	معامل الصدق التقاطعي
الصورة الأولى مع الثانية	0.000000004	0.00000002	0.0000002
الصورة الأولى مع الثالثة	0.000000001	0.000000005	0.00000007

يتضح من خلال الجدول (20) أن أقل قيمة لمعامل الصدق التقاطعي كانت للمعادلة وفق النموذج أحادي المعلمة، وهذا يشير إلى أن المعادلة وفق النموذج أحادي المعلمة كانت الأكثـر فاعلية من بين الطرق الثلاث المستخدمة. يليها المعادلة وفق النموذج ثنائي المعلمـة وأخيـراً المعادلة وفق النموذج ثلاثي المعلمة.

النتائج المتعلقة بالسؤال الثالث

وفيما يتعلق بالإجابة عن السؤال الثالث من أسئلة الدراسة والمتعلق بمقارنة فاعلية طرق المعادلة تم إعداد الصور الجديدة للاختبار والمكونة من الفقرات المطابقة لنماذج نظرية الفقـرة من صور الاختبار الأصلية. ثم تمت معادلة درجات الصور الجديدة من الاختبار بطريقتين من طرق المعادلة التي ترتكز على نظرية الاختبار، وهي:

الطريقة الأولى: المعادلة الخطية (Linear Equating):

وهنا تعتبر الدرجتان متكافئتان على صورتي الاختبار إذا كانتا تناظران العدد نفسه من الانحرافات المعيارية (فوق أو تحت) المتوسط الحسابي في مجموعة معينة مـن المفحوصـين (Angoff, 1971) وطرق المعادلة الخطية جميعها تنتج تحويلاً على الصورة التالية:

$$\ell_x(Y) = AY + B$$

حيث ℓ_x: تحويل المعادلة.

Y: علامة الاختبار التي يطبق عليها التحويل.

A,B: هما معلمتان تقدران من البيانات وتختلف قيمهما باختلاف التصميم في الدراسة.

فعند استخدام تصميم المجموعات المتكافئة – وهو المستخدم في هذه الدراسة– تكون قيم A, B كالتالي:

$$B = \overline{Y}_{eq} - \overline{Y}_x\left(\frac{\sigma_{eq}}{\sigma_x}\right) \quad ، \quad A = \left(\frac{\sigma_{eq}}{\sigma_x}\right)$$

حيث تصبح المعادلة بالصورة النهائية على الشكل التالي:

$$Y_{eq} = \frac{\sigma_{eq}}{\sigma_x}\left(Y_x - \overline{Y}_x\right) + \overline{Y}_{eq}$$

حيث:

Y_{eq}: درجات الأفراد في نموذج الاختبار المعادل إلى درجات الأفراد في نموذج الاختبار X.

σ_{eq}: الانحراف المعياري لدرجات الأفراد في نموذج الاختبار المعادل.

σ_x: الانحراف المعياري لدرجات الأفراد في نموذج الاختبارX.

Y_x: درجات الأفراد في نموذج الاختبار X.

$\overline{Y}_{eq}$: متوسط درجات الأفراد في نموذج الاختبار المعادل.

$\overline{Y}_x$: متوسط درجات الأفراد في نموذج الاختبارX.

وهذه الصيغة تناسب تصميم المجموعات المتكافئة المستخدم في هذه الدراسة. حيث أن صيغة المعادلة الخطية تختلف باختلاف التصميم المراد استخدامه في الدراسة.

تم تحويل الدرجات الخام على صور الاختبار إلى ما يناظرها على تدريج درجات الصورة الأولى، وفقاً للمعادلة التالية:

$$\ell_x(Y) = AY + B$$

حيث يبين الجدول (21) قيم المعلمتين A,B لتحويل درجات كل من الصورتين الثانية والثالثة إلى الصورة الأولى:

الجدول (21)

قيم المعلمتين A,B لتحويل درجات كل من الصورتين الجديدتين الثانية والثالثة للاختبار إلى الصورة الجديدة الأولى للاختبار

النموذج	معاملات المعادلة الخطية لكلٍ من:	A	B
أحادي المعلمة	الصورة الأولى مع الثانية	0.82	0.87
	الصورة الأولى مع الثالثة	1.08	-0.45
ثنائي المعلمة	الصورة الأولى مع الثانية	0.93	2.57
	الصورة الأولى مع الثالثة	0.96	-1.02
ثلاثي المعلمة	الصورة الأولى مع الثانية	0.94	0.18
	الصورة الأولى مع الثالثة	0.99	-1.57

وبذلك تصبح معادلات التحويل الخطي من نموذج الصورة الأولى إلى نموذج الـصورة الثانية، والأولى إلى الثالثة كما في الجدول (22):

جدول (22)

معادلات التحويل الخطي لدرجات كل من الصورتين الجديدتين الثانية والثالثة إلى الصورة الجديدة الأولى للاختبار

النموذج	المعادلة الخطية لكلٍ من:	المعادلة الخطية
أحادي المعلمة	الصورة الأولى مع الثانية	$\ell_x(Y) = 0.82Y + 0.87$
	الصورة الأولى مع الثالثة	$\ell_x(Y) = 1.08Y - 0.45$
ثنائي المعلمة	الصورة الأولى مع الثانية	$\ell_x(Y) = 0.93Y + 2.57$
	الصورة الأولى مع الثالثة	$\ell_x(Y) = 0.96Y - 1.02$
ثلاثي المعلمة	الصورة الأولى مع الثانية	$\ell_x(Y) = 0.94Y + 0.18$
	الصورة الأولى مع الثالثة	$\ell_x(Y) = 0.99Y - 1.57$

والجدول (23) يبين الدرجات المتعادلة وفق هذه المعادلات:

الجدول (23)

الدرجات المتعادلة في صور الاختبارات الجديدة الثلاث وفق المعادلة الخطية

الدرجة المشاهدة في الصورة الأولى	النموذج: أحادي المعلمة		ثنائي المعلمة		ثلاثي المعلمة	
	الدرجة المعادلة في الصورة الثانية	الدرجة المعادلة في الصورة الثالثة	الدرجة المعادلة في الصورة الثانية	الدرجة المعادلة في الصورة الثالثة	الدرجة المعادلة في الصورة الثانية	الدرجة المعادلة في الصورة الثالثة
2	2.49	1.72				
3	3.31	2.80				
4	4.14	3.88				
5	4.96	4.96				
6	5.78	6.04				
7	6.60	7.12				

الدرجة المشاهدة في الصورة الأولى	النموذج					
	أحادي المعلمة		ثنائي المعلمة		ثلاثي المعلمة	
	الدرجة المعادلة في الصورة الثانية	الدرجة المعادلة في الصورة الثالثة	الدرجة المعادلة في الصورة الثانية	الدرجة المعادلة في الصورة الثالثة	الدرجة المعادلة في الصورة الثانية	الدرجة المعادلة في الصورة الثالثة
8	7.42	8.20	9.88	6.71		
9	8.24	9.28	10.82	7.67		
10	9.06	10.36	11.75	8.63		
11	9.89	11.43	12.68	9.59		
12	10.71	12.51	13.62	10.54	11.51	10.18
13	11.53	13.59	14.55	11.50	12.45	11.18
14	12.35	14.67	15.49	12.46	13.39	12.17
15	13.17	15.75	16.42	13.42	14.33	13.16
16	13.99	16.83	17.36	14.38	15.27	14.16
17	14.81	17.91	18.29	15.33	16.20	15.15
18	15.64	18.99	19.23	16.29	17.14	16.15
19	16.46	20.07	20.16	17.25	18.08	17.14
20	17.28	21.15	21.10	18.21	19.02	18.14
21	18.10	22.23	22.03	19.17	19.96	19.13
22	18.92	23.31	22.97	20.12	20.89	20.12
23	19.74	24.39	23.90	21.08	21.83	21.12
24	20.56	25.47	24.83	22.04	22.77	22.11
25	21.39	26.55	25.77	23.00	23.71	23.11
26	22.21	27.63	26.70	23.96	24.65	24.10
27	23.03	28.71	27.64	24.92	25.59	25.10
28	23.85	29.78	28.57	25.87	26.52	26.09
29	24.67	30.86	29.51	26.83	27.46	27.09
30	25.49	31.94	30.44	27.79	28.40	28.08
31	26.32	33.02	31.38	28.75	29.34	29.07
32	27.14	34.10	32.31	29.71	30.28	30.07
33	27.96	35.18	33.25	30.66	31.22	31.06
34	28.78	36.26	34.18	31.62	32.15	32.06
35	29.60	37.34	35.12	32.58	33.09	33.05
36	30.42	38.42	36.05	33.54	34.03	34.05
37	31.24	39.50	36.98	34.50	34.97	35.04
38	32.07	40.58	37.92	35.45	35.91	36.03
39	32.89	41.66	38.85	36.41	36.84	37.03
40	33.71	42.74	39.79	37.37	37.78	38.02
41	34.53	43.82	40.72	38.33	38.72	39.02
42	35.35	44.90	41.66	39.29	39.66	40.01
43			42.59	40.25	40.60	41.01
44			43.53	41.20	41.54	42.00
45			44.46	42.16	42.47	42.99
46			45.40	43.12	43.41	43.99
47			46.33	44.08	44.35	44.98
48			47.27	45.04	45.29	45.98
49			48.20	45.99	46.23	46.97
50			49.13	46.95	47.17	47.97
51			50.07	47.91	48.10	48.96

الدرجة المشاهدة في الصورة الأولى	النموذج					
	أحادي المعلمة		ثنائي المعلمة		ثلاثي المعلمة	
	الدرجة المعادلة في الصورة الثانية	الدرجة المعادلة في الصورة الثالثة	الدرجة المعادلة في الصورة الثانية	الدرجة المعادلة في الصورة الثالثة	الدرجة المعادلة في الصورة الثانية	الدرجة المعادلة في الصورة الثالثة
52			51.00	48.87	49.04	49.96
53			51.94	49.83	49.98	50.95
54			52.87	50.78	50.92	51.94
55			53.81	51.74	51.86	52.94
56			54.74	52.70	52.79	53.93
57			55.68	53.66	53.73	54.93
58			56.61	54.62	54.67	55.92
59			57.55	55.58	55.61	56.92
60			58.48	56.53	56.55	57.91
61			59.42	57.49	57.49	58.90
62			60.35	58.45	58.42	59.90
63			61.28	59.41	59.36	60.89
64			62.22	60.37	60.30	61.89
65			63.15	61.32	61.24	62.88
66			64.09	62.28	62.18	63.88
67			65.02	63.24	63.12	64.87
68			65.96	64.20	64.05	65.87
69			66.89	65.16	64.99	66.86
70			67.83	66.11	65.93	67.85
71			68.76	67.07	66.87	68.85
72			69.70	68.03	67.81	69.84
73			70.63	68.99	68.75	70.84
74			71.56	69.95	69.68	71.83
75			72.50	70.91	70.62	72.83
76			73.43	71.86	71.56	73.82
77			74.37	72.82	72.50	74.81
78			75.30	73.78	73.44	75.81
79			76.24	74.74	74.37	76.80
80			77.17	75.70	75.31	77.80
81					76.25	78.79
82					77.19	79.79
83					78.13	80.78
84					79.07	81.77
85					80.00	82.77
86					80.94	83.76
87					81.88	84.76
88					82.82	85.75
89					83.76	86.75

يلاحظ من الجدول (23) أنه لم يتم جدولة الدرجات الخام التي قيمها أقــل مــن (2) أو التي تزيد عن (42) في صور الاختبار المكون من الفقرات المطابقة للنموذج أحادي المعلمــة؛

لأنه لم يحصل أي مفحوص على صورة الاختبار الأولى على درجة أقل من(2)، وكــذلك لــم يحصل أي مفحوص على صورة الاختبار الأولى على درجة تزيد عن (42). كما يلاحظ أنه لم يتم جدولة الدرجات الخام التي قيمها أقل من (8) أو التي تزيد عن (80) في صــور الاختبــار المكون من الفقرات المطابقة للنموذج ثنائي المعلمة؛ لأنه لم يحصل أي مفحوص على صــورة الاختبار الأولى على درجة أقل من(8)، وكذلك لم يحصل أي مفحوص على صورة الاختبــار الأولى على درجة تزيد عن (80). كذلك يلاحظ أنه لم يتم جدولة الدرجات الخام التي قيمها أقل من (12) أو التي تزيد عن (89) في صور الاختبار المكون من الفقرات المطابقة للنموذج ثلاثي المعلمة؛ لأنه لم يحصل أي مفحوص على صورة الاختبار الأولى على درجة أقــل مــن(12)، وكذلك لم يحصل أي مفحوص على صور الاختبار جميعها على درجة تزيد عن (89).

الطريقة الثانية: معادلة الاختبارات بواسطة الرتب المئينية المتــساوية (Equipercenetile Equating):

لإجراء المعادلة المئينية تم إيجاد الدرجات الخام التي تناظر الرتب المئينيــة، و يبــين الجدول (24) الدرجات الخام المتعادلة للدورات الثلاث عند الرتب المئينية من 1 إلى 99:

الجدول (24)

العلامات الخام المتعادلة في صور الاختبار الجديدة عند الرتب المئينية المختارة في المعادلة المئينية

الرتبة المئينية	أحادي المعلمة: الأولى	أحادي المعلمة: الثانية	أحادي المعلمة: الثالثة	ثنائي المعلمة: الأولى	ثنائي المعلمة: الثانية	ثنائي المعلمة: الثالثة	ثلاثي المعلمة: الأولى	ثلاثي المعلمة: الثانية	ثلاثي المعلمة: الثالثة
1	5	4.55	5	13	15	13	16	17	16
2	6	6	6	14	16	13	17.6	18	17
3	6.4	6	7	15	17	15	19	19	18
4	7	6	7	16	18	15	20	20	19
5	8	7	8	17	18	15	21	20	19
6	8	7	8	18	19	16	21	21	20
7	8	7	8	18	19	16	21.6	21	20
8	8	8	9	18	19.4	17	22	21	21

الرتبة المئينية	أحادي المعلمة - الصورة الأولى	أحادي المعلمة - الصورة الثانية	أحادي المعلمة - الصورة الثالثة	ثنائي المعلمة - الصورة الأولى	ثنائي المعلمة - الصورة الثانية	ثنائي المعلمة - الصورة الثالثة	ثلاثي المعلمة - الصورة الأولى	ثلاثي المعلمة - الصورة الثانية	ثلاثي المعلمة - الصورة الثالثة
9	9	8	9	19	20	17	22	22	21
10	9	8	9	19	20	17	23	22	21
11	9	8	9	20	21	18	23	22	22
12	9	8	10	20	21	18	24	23	22
13	9	8	10	21	21.15	19	24.4	23	23
14	9	9	10	21	22	19	25	23	23
15	10	9	10	21	22	19	25	24	23
16	10	9	10	22	22	20	26	24	24
17	10	9	11	22	22	20	26	25	24
18	10	9	11	22	22	20	26	25	25
19	10	9.45	11	23	23	21	27	25	25
20	10	10	11	23	23	21	27	25	25
21	11	10	11	24	23	21	28	26	26
22	11	10	11	24	24	22	28	26	26
23	11	10	11	24	24	22	28	26	26
24	11	10	11	25	25	22	29	27	26
25	11	10	12	25	25	22	29	27	27
26	12	10	12	25	26	23	29	28	27
27	12	10	12	25	26	23	30	28	27
28	12	11	12	26	26	23	30	28	28
29	12	11	12	26	27	23	31	28.95	28
30	12	11	12	26	27	24	31	29	29
31	12	11	12	27	27	24	31.8	29	29
32	12	11	13	27	28	24	32	30	29
33	13	11	13	27	28	24	32	30	29
34	13	11	13	27	28	25	33	30	30
35	13	11	13	28	28	25	33	31	30
36	13	11	13	28	29	25	33	31	31
37	13	11	13	28	29	26	34	31	31
38	13	12	14	29	29	26	34	32	31
39	14	12	14	29	30	26	35	32	32
40	14	12	14	30	30	27	35	32	32
41	14	12	14	30	30	27	35	33	33
42	14	12	14	30	31	27	36	33	33
43	14	12	14	31	31	28	36	34	34
44	14	12	15	31	31	29	36	34	34
45	14	13	15	32	32	29	37	35	35
46	15	13	15	32	32	29	37	35	35
47	15	13	15	32	33	30	38	35	35
48	15	13	15	32	33	30	38	36	36
49	15	13	16	33	33.95	31	38.2	36.95	36
50	15	13	16	33	34	31	39	37	37
51	16	14	16	34	34	32	39.8	37.05	38
52	16	14	16	34	35	32	40	38	38
53	16	14	17	34.4	35	33	40	39	39
54	16	14	17	35	36	33	41	39	39

الرتبة المئينية	النموذج								
	أحادي المعلمة			ثنائي المعلمة			ثلاثي المعلمة		
	الصورة			الصورة			الصورة		
	الأولى	الثانية	الثالثة	الأولى	الثانية	الثالثة	الأولى	الثانية	الثالثة
55	16	14	17	35	36	33	41	39	40
56	16	14	17	36	37	34	42	40	40
57	16	14	17	36	37	34	42	40	40
58	17	14.9	18	37	37	35	43	40	41
59	17	15	18	37	38	35	43.2	41	41
60	17	15	18	38	38	36	44	42	42
61	17	15	18	38	38	36	44	42	42
62	17	15	19	38.6	39	37	45	43	43
63	18	15.65	19	39	39	37	45	43	43
64	18	16	19	39	40	38	45.2	44	44
65	18	16	19	40	40	38	46	44	44
66	18	16	20	40	40.3	38	47	44	45
67	19	16	20	41	41	39	48	45	46
68	19	16	20	41.4	42	39	48	46	46
69	19	17	20	43	42	40	49	46	47
70	19	17	21	43	43	41	50	47	47
71	20	17	21	44	43	41	50	47	48
72	20	18	21	45	44	42	51	48	49
73	20	18	22	45	44	42	52	48	49
74	21	18	22	46	45	43	52	49	50
75	21	18	22	46	45	43	53	49.25	51
76	21	18	23	47	46	44	54	50	52
77	22	19	23	47	46.35	44	54	51	52
78	22	19	23	48	47	45	55	51	53
79	22	19	24	48	47.45	46	55	52	53
80	23	20	24	49	48	46	56	53	55
81	23	20	24	49	48	47	57	53.55	55
82	23	20	25	50	49	48	58	54	56
83	23	20	25	51	50	48	58	55	56
84	24	21	26	52	51	49	59	56	57
85	24	21	26	52	51.75	50	60	57	57
86	24	21	27	53	52.3	50	60	58	59
87	25	21	27	54	53.85	51	61	58	59
88	25	22	27	54	54	52	62	59.4	60
89	25	22	28	55	55	52	63	60	61
90	26	22	28	56	56	53	63	61	61
91	26	23	29	57	56	54	65	61.05	63
92	27	23.6	29	58	57	55	66	63	65
93	27	24	30	59	59	56	67	65	65
94	28	25	30	60	60	57	68	66	66
95	29	25	31	62	60	58	70	67	67
96	30	26	32	64	61	59	72	68	70
97	31	26	33	66	62.35	62	74	69	73
98	33.4	27	35	68.4	65.8	64	77	71.9	76
99	35.2	28.45	37	73.2	68.45	67	81.4	75	81

يلاحظ من خلال الجدول تأثر المعادلة بقيم الدرجات العليا في كل صورة؛ إذ أن المعادلة المئينية تعاني من أثر كل من سقف وأرضية الاختبار، إذ أنه بالتعريف تكون أكبر درجة ملاحظة على أحد الاختبارين معادلة لأكبر درجة ملاحظة على الاختبار الآخر، وهذا لا يحدث في المعادلة الخطية.

التحقق من معايير فاعلية المعادلة

حيث تم التحقق من دقة معادلة الاختبارات وفاعليتها في طرق المعادلة المستخدمة من خلال محكي الخطأ المعياري للمعادلة، و الصدق التقاطعي.

- محك الخطأ المعياري (Standard Error):

فقد تم إيجاد قيم الخطأ المعياري عند قيم مختارة من العلامات المعيارية لصور الاختبار المكونة من الفقرات المطابقة لنماذج نظرية الفقرة (الصورة الأولى مع الثانية، والصورة الأولى مع الثالثة)، ويبين الجدول (25) الخطأ المعياري للمعادلة الخطية الأفقية لصور الاختبارات عند علامات معيارية مختارة.

الجدول (25)
الخطأ المعياري للمعادلة الخطية للصورة الجديدة الأولى مع الصورتين الجديدتين الثانية و الثالثة عند علامات معيارية مختارة على مستوى نماذج نظرية الفقرة

Z	النموذج: أحادي المعلمة		ثنائي المعلمة		ثلاثي المعلمة	
	الخطأ المعياري لمعادلة الصورتين الأولى و الثانية	الخطأ المعياري لمعادلة الصورتين الأولى و الثالثة	الخطأ المعياري لمعادلة الصورتين الأولى والثانية	الخطأ المعياري لمعادلة الصورتين الأولى و الثالثة	الخطأ المعياري لمعادلة الصورتين الأولى والثانية	الخطأ المعياري لمعادلة الصورتين الأولى والثالثة
0.0	0.32	0.46	0.93	0.94	0.99	1.04
± 0.5	0.49	0.63	0.98	0.99	1.05	1.1
± 1.0	0.56	0.73	1.13	1.15	1.22	1.27
± 1.5	0.67	0.87	1.35	1.37	1.45	1.52
± 2.0	0.80	1.03	1.60	1.62	1.72	1.8
± 2.5	0.93	1.21	1.88	1.66	2.02	2.11
± 3.0	1.08	1.40	2.17	2.20	2.33	2.44

يلاحظ من خلال الجدول (25) أن دقة المعادلة الخطية تبلغ أكبر قيم لها عند الوسط الحسابي ($Z_x = 0$)، أي أن قيمة الخطأ المعياري يكون أقل ما يمكن، ويزداد هذا الخطأ بالابتعاد عن الوسط من أي من الجهتين ويعود ذلك إلى أخطاء المعاينة التي تؤثر على تقديرات المتوسطات الحسابية والانحرافات المعيارية المحسوبة للنماذج المختلفة (Thorandike, 1982).

وتم التحقق من دقة وفاعلية المعادلة المئينية لصور الاختبار المكونة من الفقرات المطابقة لنماذج نظرية الفقرة (الصورة الأولى مع الثانية، والصورة الأولى مع الثالثة)، باستخدام محك الخطأ المعياري، و يبين الجدول (26) قيم الخطأ المعياري لمعادلة الدورة الأولى والدورة الثانية، و الدورة الأولى والدورة الثالثة عند قيم مختارة للعلامة المعيارية z.

الجدول (26)
الخطأ المعياري للمعادلة المئينية للصورة الجديدة الأولى مع الصورتين الجديدتين الثانية و الثالثة عند علامات معيارية مختارة على مستوى نماذج نظرية الفقرة

z	أحادي المعلمة		ثنائي المعلمة		ثلاثي المعلمة	
	الخطأ المعياري لمعادلة الصورتين الأولى و الثانية	الخطأ المعياري لمعادلة الصورتين الأولى و الثالثة	الخطأ المعياري لمعادلة الصورتين الأولى والثانية	الخطأ المعياري لمعادلة الصورتين الأولى و الثالثة	الخطأ المعياري لمعادلة الصورتين الأولى والثانية	الخطأ المعياري لمعادلة الصورتين الأولى والثالثة
0.0	0.58	0.75	1.16	1.18	1.25	1.30
± 0.5	0.60	0.78	1.22	1.20	1.31	1.36
± 1.0	0.70	0.90	1.40	1.42	1.50	1.57
± 1.5	0.89	1.15	1.79	1.81	1.92	2.01
± 2.0	1.27	1.65	2.57	2.59	2.75	2.88
± 2.5	2.07	2.68	4.16	4.21	4.46	4.67
± 3.0	3.77	4.89	7.60	7.69	8.15	8.52

(النموذج)

كما يلاحظ من خلال الجدول (26) أن دقة المعادلة الخطية تبلغ أكبر قيم لها عند الوسط الحسابي ($Z_x = 0$)، أي أن قيمة الخطأ المعياري يكون أقل ما يمكن، ويزداد هذا الخطأ بالابتعاد عن الوسط من أي من الجهتين ويعود ذلك إلى أخطاء المعاينة التي تؤثر على تقديرات

المتوسطات الحسابية والانحرافات المعيارية المحسوبة للنماذج المختلفة (Thorandike, 1982).

- **محك الصدق التقاطعي (Cross Validation):**

للتحقق من دقة وفاعلية المعادلة الخطية الأفقية للفقرات المطابقة لنماذج نظرية الفقرة باستخدام محك الصدق التقاطعي، والذي يبين درجة استقرار العلامات المكافئة أو المعادلة في عينتين عشوائيتين مستقلتين، فقد تم اختيار عينة من الصورة الجديدة الأولى للاختبار، وعينتين عشوائيتين مستقلتين من الصورة الجديد الثانية للاختبار (200) فرد لكل عينة من الأفراد المطابقين لكل نموذج من نماذج نظرية الفقرة. وعينتين عشوائيتين مستقلتين من الصورة الجديد الثالثة للاختبار (200) فرد لكل عينة من الأفراد المطابقين لكل نموذج من نماذج نظرية الفقرة .

وبعد ذلك تمت معادلة الدرجات في كل من صور الاختبار الثلاث وعلى مستوى النموذج وحولت باستخدام الطريقة الخطية الأفقية من مقياس الصورة الثانية إلى الأولى، ومن مقياس الصورة الثالثة إلى الأولى، ثم حُسب معاملات الصدق التقاطعي فكانت قيمها كما في الجدول (27):

الجدول (27)
معامل الصدق التقاطعي للمعادلة الخطية للصورة الجديدة الأولى مع الصورتين الجديدتين الثانية و الثالثة على مستوى نماذج نظرية الفقرة

الصور المعادلة	النموذج		
	أحادي المعلمة	ثنائي المعلمة	ثلاثي المعلمة
الصورة الأولى مع الثانية	0.0000052	0.0000054	0.0000059
الصورة الأولى مع الثالثة	0.0000010	0.0000011	0.0000013

يلاحظ من الجدول (27) بأن معاملات الصدق التقاطعي للمعادلــة الخطيــة للــصور المطابقة للنموذج أحادي المعلمة كانت هي الأقل، وهذا يشير إلى أن المعادلة للصور المطابقــة للنموذج أحادي المعلمة كانت هي الأكثر فاعلية.

وللتحقق من دقة وفاعلية المعادلة المئينية للفقرات المطابقة لنمــاذج اســتجابة الفقــرة باستخدام محك الصدق التقاطعي، والذي يبين درجة استقرار العلامات المكافئة أو المعادلة فــي عينتين عشوائيتين مستقلتين عن بعضهما البعض، فقد تم اختيار عينة من الصورة الجديدة الأولى للاختبار، وعينتين عشوائيتين مستقلتين من الصورة الجديد الثانية للاختبار (200) فرد لكل عينة من الأفراد المطابقين لكل نموذج من نماذج استجابة الفقرة. وعينتين عشوائيتين مستقلتين مــن الصورة الجديد الثالثة للاختبار (200) فرد لكل عينة من الأفراد المطابقين لكــل نمــوذج مــن نماذج استجابة الفقرة .

ثم تمت معادلة درجات هذه العينات وفق المعادلة المئينية، ثم حُسب معامــل الــصدق التقاطعي فكانت قيمه كما في الجدول (28):

جدول (28)

معامل الصدق التقاطعي للمعادلة المئينية للصورة الجديدة الأولى مع الصورتين الجديدتين الثانية و الثالثة على مستوى نماذج نظرية الفقرة

الصور المعادلة	النموذج		
	أحادي المعلمة	ثنائي المعلمة	ثلاثي المعلمة
الصورة الأولى مع الثانية	0.00000291	0.00000320	0.00000340
الصورة الأولى مع الثالثة	0.00000076	0.00000079	0.00000111

يلاحظ من الجدول (28) بأن معاملات الصدق التقاطعي للمعادلــة المئينيــة للفقــرات المطابقة للنموذج أحادي المعلمة كانت هي الأقل، وهذا يشير إلى أن المعادلة للفقرات المطابقــة

للنموذج أحادي المعلمة كانت هي الأكثر فاعلية. وهذه النتيجة تتفق مع النتيجة التي توصلنا إليها بالنسبة لمؤشر الصدق التقاطعي للمعادلة الخطية.

كما يتضح لنا من خلال الجداول (27) ، (28) أن طريقة المعادلة المئينية كانت أكثـــر فاعلية من طريقة المعادلة الخطية في معادلة الصور المطابقة لنماذج نظرية الفقـــرة الـــثلاث؛ وذلك لأن قيم معاملات الصدق التقاطعي للمعادلة المئينية كانت أقل من قيمه للمعادلة الخطية.

كذلك نلاحظ من خلال الجداول (20)، (27)، (28) أن قيم معاملات الصدق التقـــاطعي كان لها أقل قيمة عند المعادلة وفق نماذج نظرية الفقرة تليها المعادلـــة المئينيـــة ثـــم المعادلـــة الخطية، مما يدل على أن المعادلة وفق نماذج نظرية الفقرة كانت أكثر كفاءة وفاعلية من طرق المعادلة وفق نظرية الاختبار.

الفصل الخامس

المناقشة والاستنتاجات

يتناول هذا الفصل مناقشة النتائج المتعلقة بأسئلة الدراسة، ولكل سؤال على إنفراد، مرتبة وفق ترتيب الأسئلة.

أولاً : مناقشة النتائج المتعلقة بالسؤال الأول:

كان السؤال الأول يتعلق بالخصائص السيكومترية لصور الامتحان الوطني للغة الإنجليزية في الأردن. وحيث أن الصور الثلاث للاختبار تتضمن العدد نفسه من الفقرات التي تقيس كل مهارة من المهارات: الاستيعاب السمعي (Listening Comprehension)، الاستيعاب القرائي (Reading Comprehension)، القواعد (Structure)، التعابير المكتوبة (Written Expressions) ، المفردات (Vocabulary)، وأنه تم بناء الصور وفق جدول مواصفات يمكن أن يُعد الاختبار متحقق فيه صدق البناء.

وقد تم وحسب النظرية الكلاسيكية تحليل البيانات باستخدام البرنامج الإحصائي (SPSS)، والبالغ عددها الدورة الأولى (979) مفحوصاً في الدورة الأولى، وفي الدورة الثانية (854) مفحوصاً، وفي الدورة الثالثة (899) مفحوصاً. حيث كان حجم العينة الكلي (2732) مفحوصاً.

ومن الخصائص السيكومترية للاختبار معاملات صعوبة ومعاملات تمييز فقراته، وحيث أن الحد المقبول لصعوبة الفقرة يتراوح بين (0.2–0.8) (عودة، 2003)، فقد خرجت بعض الفقرات عن هذا المدى وعددها (5) فقرات في الصورة الأولى للاختبار،(6) فقرات في

الصورة الثانية للاختبار، و(6) فقرات في الصورة الثالثة للاختبار، إلا أن المتوسطات الحسابية لمعاملات الصعوبة بقيت ضمن المدى المقبول. حيث تراوحت قيم معاملات الصعوبة للصورة الأولى (0.145-0.779) وبمتوسط حسابي (0.42)، وبين (0.181-0.751) للصورة الثانية وبمتوسط حسابي (0.413)، وبين (0.115-0.788) للصورة الثالثة وبمتوسط حسابي (0.414).

يلاحظ تقارب القيم العددية لمتوسطات معاملات الصعوبة، ولكن ذلك لا يعني بالضرورة أنه لا توجد فروق دالة إحصائياً بين متوسطات معاملات الصعوبة تُعزى لصورة الاختبار، ولذلك قام الباحث بتحويل معاملات الصعوبة إلى قيم دلتا المناظرة والتي هي درجات معيارية محولة بوسط حسابي قدره (13) وانحراف معياري قدره (4)؛ وذلك لأن قيم دلتا المناظرة تقع في مقياس فئوي، وحسب المعادلة التالية $\Delta = 4Z + 13$. ولمعرفة مدى الاختلاف بين الأوساط الحسابية لمعاملات الصعوبة للصور الثلاث للاختبار (وبافتراض أن صور الاختبار الثلاث هي صور مختلفة في الشكل وتقيس المحتوى نفسه والأهداف نفسها) فقد بين تحليل التباين الأحادي عدم وجود فروق دالة إحصائياً بين متوسطات معاملات صعوبة الفقرات للصور الثلاث عند مستوى الدلالة ($\alpha = 005$) يُعزى لصورة الاختبار. مما يعني أن فقرات النماذج الثلاث متكافئة من حيث صعوبتها وسهولتها.

كما كانت متوسطات معاملات التمييز لفقرات الصور الثلاث أكبر من (0.25)، وحسب رأي إبل (Ebel, 1972)، فإن الفقرة تمييزها أكبر من أو يساوي (0.25) تتمتع بقدرة تمييزية مقبولة إلا أن مجموعة من الفقرات في صور الاختبار كانت قيم تمييزها أقل من (0.25) وعددها في الصورة الأولى من صور الاختبار (30) فقرة، وفي الصورة الثانية من صور

الاختبار (39) فقرة، وفي الصورة الثالثة من صور الاختبار (31) فقرة. يلاحظ ارتفاع نسبة الفقرات التي كانت قدرتها التمييزية منخفضة. مما يدل على أن ضعف القدرة التمييزية لهذه الفقرات. وقد وقعت قيم معاملات التمييز ضمن المدى (−0.08 ، 0.49) للصورة الأولى وبمتوسط حسابي (0.30)، وبين (−0.14 ، 0.53) للصورة الثانية وبمتوسط حسابي (0.28)، وبين (−0.11 ، 0.59) للصورة الثالثة وبمتوسط حسابي (0.30).

يلاحظ تقارب القيم العددية لمتوسطات معاملات التمييز، ولكن ذلك لا يعني بالضرورة أنه لا توجد فروق دالة إحصائياً بين متوسطات معاملات التمييز تُعزى لصورة الاختبار، ولذلك قام الباحث بتحويل معاملات التمييز للفقرات إلى درجات فشر الزائية (Z_r) لأن إجراء هذا التحويل الرياضي على معاملات التمييز يجعل توزيع المعاينة قريباً من التوزيع الاعتدالي بغض النظر عن قيمة هذا المعامل في المجتمع أو حجم العينة، ولمعرفة مدى الاختلاف بين متوسطات معاملات التمييز في صور الاختبار الثلاث (وبافتراض أن صور الاختبار الثلاث هي صور مختلفة في الشكل وتقيس المحتوى نفسه والأهداف نفسها) تم استخدام تحليل التباين الأحادي، والذي بين عدم وجود فروق دالة إحصائياً بين متوسطات معاملات تمييز الفقرات للصور الثلاث، مما يدل على تكافؤ فقرات صور الاختبار من حيث القدرة التمييزية.

كما أظهرت النتائج أن صور الاختبار الثلاث تمتعت بثبات عالي، فقد كانت قيم معاملات ثبات الاتساق الداخلي والذي تم تقديره باستخدام معامل كرونباخ (α) لصور الاختبار الثلاث (0.918، 0.906، 0.919) على الترتيب، وعليه فإن صور الاختبار تتمتع بدقة عالية في قياس مهارات اللغة الإنجليزية للمفحوصين. وبما أن أحد شروط معادلة صور الاختبار أن تتمتع هذه الصور بالقدر نفسه من الثبات فإن هذا الشرط قد تحقق لتقارب قيم معاملات ثبات الصور الثلاث للاختبار.

وقد كانت قيمة معامل ثبات كرونباخ (α) للصورة الثالثة هي الأعلى عددياً، تليه الصورة الأولى ثم الثانية، وهذا دليل على أن متوسط معاملات الارتباط الداخلية بين فقرات الصورة الثالثة لها أعلى قيمة يليها متوسط معاملات الارتباط الداخلية بين فقرات الصورة الأولى ثم الثانية. وتنسجم هذه النتيجة مع قيم متوسطات معاملات التمييز، حيث أن الاختبار الأكثر تمييزاً هو الأكثر ثباتاً.

كذلك فقد تم حساب قيم الخطأ المعياري في القياس وقد وجد أن قيمه منخفضة ولصور الاختبار الثلاث، وهذا من شأنه التقليل مصادر الخطأ؛ حيث أنه عندما يقترن هذا الخطأ بخطأ المعادلة يصبح في هذه الحالة هنالك مصدران للخطأ، وهنا يجب أن ينصب الجهد على التقليل من نوعي الخطأ حتى تكون عملية التحويل من الدرجات الخام إلى الدرجات المتعادلة على أكبر درجة من الدقة أو أقل ما يمكن من الخطأ.

وعموماً فإن الصور الثلاث للاختبار تمتعت بخصائص سيكومترية مقبولة من حيث الصدق والثبات والصعوبة إلا أن تمييز مجموعة من فقراته كان منخفضاً، وهذه الفقرات ينصح بحذفها.

ثانياً: مناقشة النتائج المتعلقة بالسؤال الثاني والمتعلق بقيم القدرات المتعادلة في صور الاختبار الثلاث للامتحان الوطني للغة الإنجليزية في الأردن باستخدام نماذج نظرية الفقرة (أحادي المعلمة، وثنائي المعلمة، وثلاثي المعلمة)

فيما يتعلق بالإجابة عن هذا السؤال فقد تم التحقق من توافق البيانات مع افتراضات نماذج نظرية الفقرة وأهمها أحادية البعد، فقد أظهرت نتائج التحليل العاملي بطريقة المكونات الأساسية الرئيسة ثم التدوير المتعامد وجود عامل واحد يفسر وبفارق كبير عن العوامل الأخرى التي كانت قيمة الجذر الكامن لها أكبر من واحد ولصور الاختبار الثلاث، حيث كانت نسبة

الجذر الكامن للعامل الأول إلى الجذر الكامن للعامل الثاني تزيد عن (2)، مما يُعطي دليلاً كافياً على أحادية البعد لصور الاختبار الثلاث.

أما بالنسبة لافتراض الاستقلال الموضعي فيرى هامبلتون وسوامنثان (Hambleton and Swaminathan, 1985) وبيجار (Bejar, 1980) أن هذا الافتراض مكافئ لفرض أحادية البعد للسمة الكامنة. فإذا لم تكن الفقرات مستقلة إحصائياً (غير مترابطة) عند مستوى قدرة معين فهذا يعني بأن العلامات المتوقعة لبعض المفحوصين على الاختبار ستكون أعلى مقارنة ببعض المفحوصين ممن هم بمستوى القدرة نفسه وبالتالي لا بد أن تكون هناك أكثر من قدرة لتفسير أداء المفحوصين على الاختبار وهذا انتهاك للافتراض القائل: إن الفقرات أحادية البعد. ومن المهم هنا الإشارة إلى أن افتراض الاستقلال الموضعي لا يتضمن أو يعني أن الفقرات غير مترابطة عبر المجموعة الكلية للمفحوصين ولكن العلامات غير المترابطة عند مستوى قدرة معين. ولذا سمي مستقل محلياً أو موضعياً (Locally Independent). وبناءاً عليه فإن هذا الافتراض متحقق في بيانات الدراسة.

وفيما يتعلق بدرجة مطابقة الاستجابات لفقرات الاختبار مع النموذج أحادي المعلمة من نماذج نظرية الفقرة فقد قام الباحث بتحليل استجابات (979) فرداً في الصورة الأولى من الامتحان عن (100) فقرة، و(854) فرداً في الصورة الثانية من الامتحان عن (100) فقرة، و (899) فرداً في الصورة الثالثة من الامتحان عن (100) فقرة باستخدام برنامج (BILOG) والمصمم لتحليل البيانات وفقاً لنماذج نظرية الفقرة، ولكن البرنامج حلل استجابات (938) فرداً بعد أن حذف (41) فرداً في الصورة الأولى من الامتحان ، وحلل استجابات (811) فرداً بعد أن حذف (43) فرداً في الصورة الثانية من الامتحان ، وحلل استجابات (867) فرداً بعد أن حذف (32) فرداً في الصورة الثالثة من الامتحان. بينما عند مطابقة الاستجابات لفقرات الاختبار مع

النموذج ثنائي المعلمة من نماذج نظرية الفقرة فقد حلل البرنامج استجابات (946) فرداً بعد أن حذف (33) فرداً في الصورة الأولى من الامتحان ، وحلل استجابات (818) فرداً بعد أن حذف (36) فرداً في الصورة الثانية من الامتحان ، وحلل استجابات (869) فرداً بعد أن حذف (30) فرداً في الصورة الثالثة من الامتحان.و عند مطابقة الاستجابات لفقرات الاختبار مع النموذج ثلاثي المعلمة من نماذج نظرية الفقرة فقد حلل استجابات (944) فرداً بعد أن حذف (35) فرداً في الدورة الأولى من الامتحان ، وحلل استجابات (818) فرداً بعد أن حذف (36) فرداً في الدورة الثانية من الامتحان ، وحلل استجابات (867) فرداً بعد أن حذف (32) فرداً في الدورة الثالثة من الامتحان. وهؤلاء الأفراد الذين تم حذفهم لم تتطابق استجاباتهم الملاحظة مع الاستجابات المتوقعة تبعاً لقدراتهم، حيث أن استجاباتهم تبتعد عن توقعات النموذج. وأعيد التحليل للمرة الثانية بعد حذف الأفراد غير المطابقين لكل لنموذج من نماذج نظرية الفقرة لحذف الفقرات التي لا تتطابق مع توقعات النموذج حيث أن حذف الأفراد من التحليل يحدث مرة واحدة ولا يُعاد الحذف حتى لو أظهر التحليل أفراداً جدداً غير مطابقين من التحليل التالي. وفي ضوء نتائج التحليل تبين وجود (58) فقرة غير متطابقة مع توقعات النموذج أحادي المعلمة في الصورة الأولى، و (65) في الصورة الثانية، و (55) في الصورة الثالثة. وتبين وجود (13) فقرة غير متطابقة مع توقعات النموذج ثنائي المعلمة في الصورة الأولى، و (14) في الصورة الثانية، و (21) في الصورة الثالثة. وتبين وجود (2) فقرة غير متطابقة مع توقعات النموذج ثلاثي المعلمة في الصورة الأولى، و (6) في الصورة الثانية، و (5) في الصورة الثالثة.حيث تكون الاستجابات الملاحظة عنها تبتعد عن تلك التي يتوقعها النموذج بشكل جوهري. وأعيد التحليل مرة ثالثة بعد حذف الفقرات، حيث تكونت الصورة النهائية للاختبار وتم الحصول على النتائج النهائية المتحررة من صعوبة الفقرات وقدرات الأفراد.

وأظهرت النتائج أن عدد الفقرات التي تم حذفها، وعدد الأفراد الذين تم حذف استجاباتهم عند المطابقة كانت الأكثر عند النموذج أحادي المعلمة، يليها عدد الفقرات المحذوفة عدد الأفراد الذين تم حذف استجاباتهم عند استخدام النموذج ثنائي المعلمة وأخيراً عدد الفقرات التي تم حذفها عدد الأفراد الذين تم حذف استجاباتهم عند استخدام النموذج ثلاثي المعلمة، وهذا أمر منطقي لأن النموذج أحادي المعلمة يفترض اختلاف الفقرات في معلمة الصعوبة فقط، أما النمـوذج ثنـائي المعلمة فيفترض اختلاف الفقرات في معلمتين هما: الصعوبة والتمييز، أمـا النمـوذج ثلاثـي المعلمة فيفترض اختلاف الفقرات في ثلاث معالم هي: الصعوبة والتمييز، والتخمين. فـالنموذج أحادي المعلمة هو الأكثر تشدداً. ولمعرفة مدى الاختلاف في نسب الفقرات المطابقـة ونـسب الأفراد الذين تم حذف استجاباتهم لكل نموذج من نماذج الفقرة فقد بينت نتيجة التحليل عدم وجود فروق دالة إحصائياً بين نسب الفقرات المطابقة لنماذج الفقرة. بينما بينت نتيجة التحليل وجـود فروق دالة إحصائياً بين نسب الأفراد الذين طابقت استجاباتهم لنماذج الفقرة. **مما يؤيـد فكـرة المعادلة.**

وعند تقدير قدرات الأفراد وفق نماذج نظرية الفقرة الثلاث لوحظ ارتفاع قيم الأخطـاء المعيارية لقدرات الأفراد الذين حصلوا على أقل العلامات أو أعلاها، بينما تقل كلما اتجهنا نحو مركز التوزيع أي متوسط تقديرات القدرات، وهذا يؤكد دقة تحديد مواقع الأفراد على متـصل السمة.

وكان أعلى متوسط لتقدير القدرة للأفراد وفق النموذج أحادي المعلمـة فـي الـصورة الأولى للاختبار وبلغ (0.099) لوجيت، وهي موجبة وأعلى من متوسـط الـصعوبة للفقـرات والبالغ صفر لوجيت وبالتالي فإن المتوقع هو تزايد احتمال الإجابة الصحيحة علـى الفقـرات بتزايد مستوى قدرات الأفراد وهذا يعني اتساق معطيات الفقرات مع توقعات النموذج أي حسن

مطابقة الفقرات للنموذج. أما أدنى متوسط لتقدير القدرة للأفراد كان في الصورة الثانية وبلغ (-0.086) لوجيت، يليه الصورة الثالثة وبلغ (0.049-)، وهي سالبة وأقل من متوسط الصعوبة للفقرات والبالغ صفر لوجيت وبالتالي فإن المتوقع هو أن يقل احتمال الإجابة الصحيحة على الفقرات بنقصان مستوى قدرات. **ويُعد ذلك مؤشراً على عدم تكافؤ صور الاختبار الثلاث. وهو مبرر لإجراء عملية المعادلة.**

ولما كان أحد شروط معادلة صور الاختبار تحقيق مطلب العدالة (Equity)؛ أي أن تكون خصائص التوزيع التكراري المشروط للدرجات عند مستوى معين من القدرة (θ) للاختبار Y بعد تحويل الدرجات هي نفسها للتوزيع التكراري المشروط للدرجات عند المستوى نفسه من القدرة (θ) للاختبار X أي أن $f(Y/\theta) = f(X/\theta)$. فقد وجد أن هذا الشرط قد تحقق في هذه الدراسة؛ حيث أظهرت نتائج تحليل التباين عدم وجود فروق دالة إحصائياً بين متوسطات القدرة المقدرة (θ) المتعادلة لصور الاختبار الثلاث وعلى مستوى نماذج نظرية الفقرة.

<u>ثالثاً</u>: مناقشة النتائج المتعلقة بالسؤال الثالث والمتعلق بكفاءة ودقة طرق المعادلة

وحيث أن أحد شروط المعادلة أن يكون تحويل نظام وحدات القياس مستقلاً عن عينة الأفراد التي استمدت منها البيانات المستخدمة في إجراء هذا التحويل فقد تم بناء صور جديدة من الاختبار مكونة من الفقرات المطابقة لنماذج نظرية الفقرة من فقرات الصور الأصلية واستخدام عينة الأفراد المطابقين لنماذج نظرية الفقرة من عينة الأفراد الأصلية؛ وذلك لأن نماذج نظرية الفقرة تتميز بأنها متحررة من عينة الأفراد (Person Free). وهكذا يكون هذا الشرط قد تحقق في دراستنا.

لوحظ أن المعادلة الخطية للصور المكونة من الفقرات المطابقة للنموذج أحادي المعلمة كانت الأكثر فاعلية وذلك لأن قيم الخطأ المعياري لها كانت هي الأقل من بين قيم الخطأ المعياري للمعادلة وفق النماذج الثلاث. بينما كانت المعادلة الخطية للصور المكونة من الفقرات المطابقة للنموذج ثلاثي المعلمة هي الأقل فاعلية وذلك لأن قيم الخطأ المعياري لها كانت الأكبر من بين قيم الخطأ المعياري للمعادلة وفق النماذج الثلاث.

كما وجد المعادلة المئينية للصور المكونة من الفقرات المطابقة للنموذج أحادي المعلمة كانت الأكثر دقة وذلك لأن قيم الخطأ المعياري لهذه المعادلة كانت هي الأقل، وأن المعادلة المئينية للصور المكونة من الفقرات المطابقة للنموذج ثلاثي المعلمة كانت الأقل دقة وذلك لأن قيم الخطأ المعياري لهذه المعادلة كانت هي الأكبر. وهذا يوافق ما توصلنا إليه باستخدام المعادلة الخطية.

كما وجد أن طريقة المعادلة الخطية كانت أكثر فاعلية من طريقة المعادلة المئينية في معادلة الصور المكونة من الفقرات المطابقة لنماذج استجابة الفقرة الثلاث؛ وذلك لأن قيم الخطأ المعياري للمعادلة الخطية كانت أقل من قيمه للمعادلة المئينية.

كما كانت معاملات الصدق التقاطعي للمعادلة الخطية للصور المكونة من الفقرات المطابقة للنموذج أحادي المعلمة كانت هي الأقل، وهذا يشير إلى أن المعادلة للصور المكونة من الفقرات المطابقة للنموذج أحادي المعلمة كانت هي الأكثر فاعلية، بينما كانت معاملات الصدق التقاطعي للمعادلة الخطية للصور المكونة من الفقرات المطابقة للنموذج ثلاثي المعلمة هي الأكبر، مما يدل على أن المعادلة للصور المكونة من الفقرات المطابقة للنموذج ثلاثي المعلمة كانت هي الأقل فاعلية.

كما لوحظ بأن معاملات الصدق التقاطعي للمعادلة المئينية للفقرات المطابقة للنموذج أحادي المعلمة كانت هي الأقل، وهذا يشير إلى أن المعادلة للفقرات المطابقة للنموذج أحادي المعلمة كانت هي الأكثر فاعلية. وهذه النتيجة تتفق مع النتيجة التي توصلنا إليها بالنسبة لمؤشر الصدق التقاطعي للمعادلة الخطية.

كما واتضح من خلال التحليل أن طريقة المعادلة المئينية كانت أكثر فاعلية من طريقة المعادلة الخطية في معادلة الصور المطابقة لنماذج استجابة الفقرة الثلاث؛ وذلك لأن قيم معاملات الصدق التقاطعي للمعادلة المئينية كانت أقل من قيمه للمعادلة الخطية.

كذلك لوحظ من خلال التحليل أن قيم معاملات الصدق التقاطعي كان لها أقل قيمة عند المعادلة وفق نماذج الفقرة تليها المعادلة المئينية ثم المعادلة الخطية، مما يدل على أن المعادلة وفق نماذج الفقرة كانت أكثر كفاءة وفاعلية من طرق المعادلة الكلاسيكية. وهذا يتفق مع النتيجة التي توصلت إليها دراسات كل من كولن و وتني (Kolen & Whitney, 1982)، أيوب (1994)، الشريفين (2003)، ودراسة هانسون وكيم (Hanson & Kim, 2002). التي أشارت جميعها إلى فاعلية نماذج نظرية الفقرة. ويخالف ما توصلت إليه دراسة سكاجز وليستر (Skagges & Lissitiz, 1986) والتي أظهرت أن المعادلة المئينية هي الأقوى والأكثر فاعلية، ودراسة بيترسون وكوك وستوكنج (Peterson, Cook &Stocking, 1983) التي أظهرت فاعلية المعادلة الخطية وتفوقها، كما تعارضت مع دراسة فاندرليندن Vander (Linden, 2006 التي أشارت إلى فاعلية المعادلة المئينية وأنها تعطي خطأ أقل للمعادلة.

التوصيات والاقتراحات:

من خلال نتائج الدراسة يوصي الباحث بما يلي :

- أظهرت نتائج الدراسة أن مجموعة من فقـرات صـور الاختبـار الـثلاث لا تتمتـع بخصائص سيكومترية جيدة، ولذلك يوصي الباحث بعدم استخدام فقراتها في الـدورات التالية، وخصوصاً الفقرات التي كانت قيم تمييزها منخفضة أو سالبة.
- يوصي الباحث بإجراء دراسات حول نماذج أخرى غير المستخدمة في هذه الدراسة من نماذج الامتحان الوطني للغة الإنجليزية في الأردن.
- أظهرت طرق نظرية الفقرة في المعادلة فاعلية عالية من خلال استخدام محك الـصدق التقاطعي، ولذلك يوصي الباحث بمعادلة صور الاختبارات باستخدامها.
- يوصي الباحث بمعادلة صور الاختبارات التي تتصف بالأهمية لشرائح المجتمع، مثـل اختبارات الكفاءة لطلبة السنة الأولى في الجامعات الأردنية. ومن ثم اتخاذ القرارات بناءً على نتائج المعادلة وخصوصاً إذا كانت نماذج هذه الاختبارات غير متكافئة.
- يوصي الباحث بدراسة فقرات صور الامتحان الوطني للغة الإنجليزية في الأردن مـن نواحي أخرى غير معادلة درجات صوره، مثل الأداء التفاضلي لفقراته.
- يقترح الباحث تكوين بنك أسئلة مكون من فقـرات صـور الامتحـان الـوطني للغـة الإنجليزية التي تتميز بمعالم صعوبة وتمييز وتخمين مقبولـة وفـق نظريـة الفقـرة؛ لاستخدام هذه الصور وتطبيقها في الدورات التالية.

المراجع:

المراجع العربية

الحواري، أروى. (2008). **الخصائص السيكومترية لصور مختارة من اختبارات الرخصة الدولية لقيادة الحاسوب (ICDL) في الأردن ومعادلة درجاتها**، رسالة ماجستير غير منشورة، جامعة اليرموك، اربد، الأردن.

الشريفين، نضال. (2003). **مدى تحقق معايير الفاعلية في معادلة اختبارين أحدهما ثنائي التدريج والآخر متعدد التدريج وفق نماذج النظرية الكلاسيكية والنظرية الحديثة في القياس.** رسالة دكتوراه غير منشورة، جامعة عمان العربية للدراسات العليا، عمان، الأردن.

الشريفين، نضال. (2009). معادلة درجات نماذج مختلفة من اختبار الكفاءة اللغوية في اللغة الإنجليزية لدى طلبة جامعة اليرموك. **مجلة جامعة أم القرى العربية للعلوم التربوية والنفسية**، جامعة أم القرى، (2)، 11-62.

أيوب، حسين. (1994). **المقارنة بين أربع طرق للمعادلة عندما يكون التصميم من مجموعات متكافئة وغير متكافئة.** رسالة دكتوراه غير منشورة، الجامعة الأردنية، عمان، الأردن.

الدوسري، راشد. (2001). معادلة الاختبارات مفهومها، وطرقها، ومشكلات تطبيقها. **مجلة العلوم التربوية والنفسية.** جامعة البحرين، كلية التربية، 2(4)، 107-139.

الصمادي، مروان. (2005). **فاعلية طرق تصحيح اختبار الصواب-الخطأ المتعدد وتأثيرها على دقة معادلـــة الاختبار باستخدام نماذج النظرية الحديثة في القياس**.رسالة دكتوراة غيرمنشورة، جامعة اليرموك، اربد، الأردن.

عبد الهادي، نبيل. (2001). **القياس والتقويم التربوي، واستخدامه في مجال التدريس الصفي**، ط2، عمان، الأردن: دار وائل للنشر.

علام، صلاح الدين. (2001). **الاختبارات التشخيصية مرجعية المحك في المجالات التربويـــة والنفسية والتدريسية**،ط1، دار الفكر العربي. القاهرة.

علام، صلاح الدين. (2005). **نماذج الاستجابة للمفردة الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي**،ط1، دار الفكر العربي. القاهرة.

عوده، أحمد. (2010) **القياس والتقويم في العملية التدريسية**، دار الأمل، اربد، الأردن.

المراجع الأجنبية

Al banese, M. A & ,. Forsyth, R. (1984). The one- tow and modified two parameter latent trait models: An empirical study of relative fit. **Educational and psychological Measurement**, 44, 229-246.

Allen, M. & Yen, W. (1979). **Introduction to Measurement Theory**. California: Cole Publishing Company.

Angoff, W. (1971). Scales, norms, and equivalent scores. In Thorndike, R. L. (Editor) **Educational Measurement**, Washington D. C., American Council in Education.

Cook, L. & Eignor, D. (1991) IRT equating methods **Educational Measurement: Issues and practice, 11**(4)**, 37**.

Crocker, L. & Algina, J. (1986) **Introduction to Classical and Modern Test Theory** New York: Holt, Rinehart & Winston.

Dorans, N. (1990) Equating Methods and Sampling. **Applied Measurement in education**, **3**(1), 3-18.

Dorans, J., Moses, P. & Eignor, R. (2010) **Principles** and **practices** of **Test Score Equating**.(Ed). By Powers D. Educational Testing Services (ETS), Princeton, New Jersey.

Ebel, R. (1972). Essential of Educational Measurement. New Jersey: Printice-Hall, INC. Englewood Cliffs.

Glowacki, M. (1991) **An analysis of test equating models for the Alabama high school graduation examination**. Paper presented at the annual meeting of Mid-South Educational Research Association Lexington. (ERIC Documents reproduction service number ED320720).

Hambleton, R. & Swaminathan, H. (1985). **Item Response Theory Principles and Applications** . kluwer, Nijhoff publishing. Boston, 1-4.

Hambleton, R., Swaminathan, H. & Rogers, H. (1991). **Fundamentals of Item Response Theory**. Newbery Park: Sage Publications.

Hanson, B. & Kim, S. (2002). Test equating under the multiple-choice model. **Applied**. **Educational and psychological Measurement**, 26(3), 255-270.

Harris, D. & Kolen, M. (1990). Effect of examinee group on equating relationships. **Applied psychological Measurement 10,** 35-43.

Hills, J. Subhiyh, R. & Hirsch, T. (1988). Equating minimum Competency tests: comparison of instructional methods. **Journal of Education Measurement. 25**, 221-232.

Holland, P. & Rubin, D. (1982) **Test Equating**. N.Y.: Academic Press.

Kim, H. & Kolen, M. (2010) The Effect of Repeaters on Squating, **Applied Measurement in Education**,(23), 242-265.

Kim, D., Seung, W., Lee, G. & Kooghyang, R. (2008) A Comparison of the Common-Item and Random-Groups Equating Designs Using Empirical Data, **International Journal of Selection and Assessment**,(16), 83-92.

Kolen, M. (1981). Comparison of traditional and item response theory of equating tests. **Journal of Educational Measurement. 18**(1), 1-11.

Kolen, M. & Brennan, R. (1995). **Test Equating :Methods and Practices**. N.Y.:Springer.

Kolen, M. & Whitney, D. (1982). Comparison of four procedures for equating the test of general educational development. **Journal of Educational Measurement,9**(4), 279-293.

Lee, G, Hill, M., Kolen, M. J. & , .Frisbie, D. A. (2001). Comparison of dichotomous and polytomous item response models in equating scores from test composed of test lets. **Applied psychological Measurement, 25** (4), 357-372.

Livingston Samuel A. (2004) **Equating Test Scores (Without IRT)** (online). 12/5/2013. www.ets.org/Media/Research/pdf/LIVINGSTON.pdf.

Livingston, S., Dorans, N. & Wright, N. (1990). What combination of sampling and equating methods works best? **Applied Measurement in Education, 3**(1), 73-96.

Lord, F . (1977). Practical applications of item characteristics curve Theory. **Journal of Educational Measurement. 14**, 117-138.

Lord, F. (1989). **Applications of Item Response Theory to Practical Testing Proplems**. Hillsdale, N.J. : Lawrence Erlbaum.

Peterson, N., Cook, L. & Socking, M. (1983). IRT versus conventional equating methods: A comparative study of scale stability. **Journal of Educational Statistics**,8, 137-156.

Peterson, N., Kolen, M. & Hoover, H. (1989). Scaling norming and equating .**Educational Measurement**, (Ed). By Linn, R. L.: 3rd ed., Washington, D.C.: American Council on Education, 241-262.

Reise, S. & Yu, J. (1990). Parameter recovery in the graded response using MULTILOG. **Journal of Education Measurement. 27** (2), 33-44.

Skaggs, G., & Lissitz, R. (1986). IRT test equating relevant issues and a review of recent research. **Review of Educational Research**. 56 (4), 495 –529.

Thorandike, R. (1982). **Applied Psychometrics**, Houghton Mifflin Company, Boston.

Vander Linden, W. (2006). Equating Scorees from adaptive to linear tests, **Applied Psychological Measurement**, 30(6), 493-508.

Wright, B. & Masters, G., (1982). **Rating Scale Analysis. Rash Measurement.** MESA Press, kimbark Avenue, Chicago, 59.

ABSTRACT

Olimat, Mahdi Mohammad Fahad. (2013). Equating Scores of Selected Forms of National Examination for English Language in Jordan Using Test Theory and Item Theory. (supervisor: Professor Ahmad Audeh).

This study aimed to equate three selected forms of the National Examination for English Language in Jordan using test theory and item theory. To achieve this aim, the data associated of the three forms was taken from the Ministry of High Education and Scientific Research, which was applied on (979, 854, and 899) examinees from a population of (7042) examinees.

The psychometric properties and the IT (Item Theory) assumptions of the three forms were verified, then get persons and items fit IT models in each test form. The new forms of test which consists of the items that met IT models were equated using the linear and equipercintile horizontal methods based on test theory. Two criteria were used to ascertain the efficiency of equating; these were cross validation coefficient and standard error of equating.

The following results were found: The Internal-Consistency coefficients for the three original forms were ((0.92, (0.91 (0.92). There is no statistical difference between the means and variances of the scores of the three forms. Also the analysis shows that there is no statistical differences between means of difficulties and discriminations of the three forms. The three original forms met the IT assumptions. There is no statistical difference between the ratio of items which fit the IT models due to the form, but there is statistical difference between the ratio of persons who fit the IT models due to the form. No significant differences between the means of estimated abilities due to the form in each of the three item

response models. The most efficient equating method was the one parameter model. The linear and equipercintile equation for the items which fit the one parameter model was the most efficient.

(Key Words: National Examination, English Language, Test Theory, Equating, Item Theory (IT)).